History

The journal *Rendiconti dell'Istituto di Matematica dell'Università di Trieste* was founded in 1969 with the aim of publishing original research articles in all fields of mathematics. The first director of the journal was Arno Predonzan, subsequent directors were Graziano Gentili, Enzo Mitidieri and Bruno Zimmermann.

Rendiconti dell'Istituto di Matematica dell'Università di Trieste has been the first Italian mathematical journal to be published also on-line. The access to the electronic version of the journal is free. All articles are available on-line.

In 2008 the *Dipartimento di Matematica e Informatica*, the owner of the journal, decided to renew it. In particular, a new Editorial Board was formed, and a group of four Managing Editors was selected. The name of the journal however remained unchanged; just the subtitle *An International Journal of Mathematics* was added. Indeed, the opinion of the whole department was to maintain this name, not to give the impression, if changing it, that a further new journal was being launched.

Website Address: http://www.dmi.units.it/~rimut/

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 1–9

On the Solvability Conditions for a Linearized Cahn-Hilliard Equation

Vitaly Volpert and Vitali Vougalter

Abstract. *We derive solvability conditions in $H^4(\mathbb{R}^3)$ for a fourth order partial differential equation which is the linearized Cahn-Hilliard problem using the results obtained for a Schrödinger type operator without Fredholm property in our preceding work [17].*

Keywords: Solvability Conditions, Non-Fredholm Operators, Sobolev Spaces
MS Classification 2010: 35J10, 35P10, 31A30

1. Introduction

Consider a binary mixture and denote by c its composition, that is the fraction of one of its two components. Then the evolution of the composition is described by the Cahn-Hilliard equation (see, e.g., [1, 11]):

$$\frac{\partial c}{\partial t} = M\Delta\left(\frac{d\phi}{dc} - K\Delta c\right), \tag{1}$$

where M and K are some constants and ϕ is the free energy density. From the Flory-Huggins solution theory it follows that

$$\frac{d\phi}{dc} = k_1 + k_2 c + k_3 T(\ln c - \ln(1-c)),$$

k_i, $i = 1, 2, 3$, are some thermodynamical constants and T is the temperature (see, e.g., [8]). We note that the constants k_1, k_2 and K characterize interaction of components in the binary medium. They can be positive or negative. If the components are identical, that is the medium is not in fact binary, they are equal to zero. In this case, equation (1) is reduced to the diffusion equation.

Denote the right-hand side of the last equality by $F(T, c)$. If the variation of the composition is small, then we can linearize it around some constant $c = c_0$:

$$F(T, c) \approx k_1 + k_2 c + k_3 T(\alpha + \beta(c - c_0)).$$

where $\alpha = \ln(c_0/(1-c_0))$ and $\beta = 1/(c_0(1-c_0))$. Substituting the expression for $F(T,c)$ into (1), we obtain

$$\frac{\partial u}{\partial t} = M\Delta\left(k_1 + k_2 c_0 + \alpha k_3 T + (k_2 + k_3\beta T)u - K\Delta u\right), \tag{2}$$

where $u = c - c_0$.

The existence, stability and some properties of solutions of the Cahn-Hilliard equation have been studied extensively in recent years (see, e.g., [3, 6, 11]). In this work we investigate the existence of stationary solutions of equation (2), which we write as

$$\Delta(\Delta u + V(x)u + au) = f(x), \tag{3}$$

where

$$V(x) = -\frac{k_3\beta T_0(x)}{K}, \quad f(x) = \frac{\alpha k_3}{K}\Delta T_0(x) + g(x), \quad a = -\frac{k_2 + k_3\beta T_\infty}{K}.$$

Here we use the representation $T(x) = T_\infty + T_0(x)$, where T_∞ denotes the value of the temperature at infinity and $T_0(x)$ decays as $|x| \to \infty$; $g(x)$ is a source term.

Thus, from the physical point of view, we study the existence of stationary composition distributions depending on the temperature distribution, which enters both in the coefficient of the equation and in the right-hand side. If the temperature distribution is constant, that is $T_0(x) \equiv 0$, then we obtain a homogeneous equation with constant coefficients. It can have either only trivial solution, in which case the composition distribution is also constant, $c \equiv c_0$, or, if the spectrum contains the origin, a nonzero eigenfunction. This case corresponds to the phase separation.

In this work we study the case of a nonuniform temperature distribution, such that $T_0(x)$ does not vanish identically. We will formulate the conditions of the existence of the solution. If a solution does not exist, then this can signify that the assumption about small variation of the composition is not applicable. Instability of the homogeneous in space solution results in phase separation with strong composition gradients.

From the mathematical point of view, we consider a linear elliptic equation of the fourth order in $\mathbb{R}^3$. There are two principally different cases. If the essential spectrum of the corresponding elliptic operator does not contain the origin, then the operator satisfies the Fredholm property, its image is closed and equation (3) is solvable if and only if $f(x)$ is orthogonal to all solutions of the homogeneous adjoint equation. The essential spectrum can be determined through limiting operators [15]. If the coefficients of the operator have limits at infinity, the essential spectrum can be easily found by means of Fourier transform (see below). If it contains the origin, then the operator does not satisfy the Fredholm property and the Fredholm alternative is not applicable.

In spite of apparent simplicity of equation (3), its solvability conditions in the non-Fredholm case are not known. In the case of the second order equation, solvability conditions were recently obtained in our previous works [17]–[20]. In this work we will apply these results to study the fourth order equation.

Let us assume that the potential $V(x)$ is a smooth function vanishing at infinity. The precise conditions on it will be specified below. The function $f(x)$ belongs to the appropriate weighted Hölder space, which will imply its square integrability, and a is a nonnegative constant. We will study this equation in $\mathbb{R}^3$.

The operator

$$Lu = \Delta(\Delta u + V(x)u + au)$$

considered as acting from $H^4(\mathbb{R}^3)$ into $L^2(\mathbb{R}^3)$ (or in the corresponding Hölder spaces) does not satisfy the Fredholm property. Indeed, since $V(x)$ vanishes at infinity, then the essential spectrum of this operator is the set of all complex λ for which the equation

$$\Delta(\Delta u + au) = \lambda u$$

has a nonzero bounded solution. Applying the Fourier transform, we obtain

$$\lambda = -\xi^2(a - \xi^2), \quad \xi \in \mathbb{R}^3.$$

Hence the essential spectrum contains the origin. Consequently, the operator does not satisfy the Fredholm property, and solvability conditions of equation (3) are not known. We will obtain solvability conditions for this equation using the method developed in our previous papers [17]–[20]. This method is based on spectral decomposition of self-adjoint operators.

Obviously, the problem above can be conveniently rewritten in the equivalent form of the system of two second order equations

$$\begin{cases} -\Delta v = f(x), \\ -\Delta u - V(x)u - au = v(x) \end{cases} \tag{4}$$

in which the first one has an explicit solution due to the fast rate of decay of its right side stated in Assumption 3, namely

$$v_0(x) := \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{f(y)}{|x-y|}dy \tag{5}$$

with properties established in Lemma A1 of the Appendix. Note that both equations of the system above involve second order differential operators without Fredholm property. Their essential spectra are $\sigma_{ess}(-\Delta) = [0, \infty)$ and $\sigma_{ess}(-\Delta - V(x) - a) = [-a, \infty)$ for $V(x) \to 0$ at infinity (see, e.g., [9]), such that neither of the operators has a finite dimensional isolated kernel. Solvability conditions for operators of that kind have been studied extensively in recent works for a single Schrödinger type operator (see [17]), sums of second order differential operators (see [18]), the Laplacian operator with the drift term (see [19]). Non Fredholm operators arise as well while studying the existence and stability of stationary and travelling wave solutions of certain reaction-diffusion equations (see, e.g., [5, 7, 16]). For the second equation in system (4) we introduce the corresponding homogeneous problem

$$-\Delta w - V(x)w - aw = 0. \tag{6}$$

We make the following technical assumptions on the scalar potential and the right side of equation (3). Note that the first one contains conditions on $V(x)$ analogous to those stated in Assumption 1.1 of [17] (see also [18, 19]) with the slight difference that the precise rate of decay is assumed not a.e. as before but pointwise since in the present work the potential function is considered to be smooth.

Assumption 1. *The potential function $V(x) : \mathbb{R}^3 \to \mathbb{R}$ satisfies the estimate*

$$|V(x)| \le \frac{C}{1 + |x|^{3.5+\delta}}$$

with some $\delta > 0$ and $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ such that

$$4^{\frac{1}{9}} \frac{9}{8} (4\pi)^{-\frac{2}{3}} \|V\|_{L^\infty(\mathbb{R}^3)}^{\frac{1}{9}} \|V\|_{L^{\frac{4}{3}}(\mathbb{R}^3)}^{\frac{8}{9}} < 1 \quad \textit{and} \quad \sqrt{c_{HLS}} \|V\|_{L^{\frac{3}{2}}(\mathbb{R}^3)} < 4\pi.$$

Here and further down C stands for a finite positive constant and c_{HLS} given on p.98 of [12] is the constant in the Hardy-Littlewood-Sobolev inequality

$$\Big| \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{f_1(x) f_1(y)}{|x-y|^2} dx dy \Big| \le c_{HLS} \|f_1\|_{L^{\frac{3}{2}}(\mathbb{R}^3)}^2, \quad f_1 \in L^{\frac{3}{2}}(\mathbb{R}^3).$$

Here and below the norm of a function $f_1 \in L^p(\mathbb{R}^3)$, $1 \le p \le \infty$ is denoted as $\|f_1\|_{L^p(\mathbb{R}^3)}$.

Assumption 2. $\Delta V \in L^2(\mathbb{R}^3)$ *and* $\nabla V \in L^\infty(\mathbb{R}^3)$.

We will use the notation

$$(f_1(x), f_2(x))_{L^2(\mathbb{R}^3)} := \int_{\mathbb{R}^3} f_1(x) \bar{f}_2(x) dx,$$

with a slight abuse of notations when these functions are not square integrable, like for instance some of those used in the Assumption 3 below. Let us introduce the auxiliary quantity

$$\rho(x) := (1 + |x|^2)^{\frac{1}{2}}, \ x = (x_1, x_2, x_3) \in \mathbb{R}^3 \tag{7}$$

and the space $C_a^{\mu}(\mathbb{R}^3)$, where a is a real number and $0 < \mu < 1$ consisting of all functions u for which

$$u\rho^a \in C^{\mu}(\mathbb{R}^3).$$

Here $C^{\mu}(\mathbb{R}^3)$ stands for the Hölder space such that the norm on $C_a^{\mu}(\mathbb{R}^3)$ is defined as

$$\|u\|_{C_a^{\mu}(\mathbb{R}^3)} := \sup_{x\in\mathbb{R}^3}|\rho^a(x)u(x)| + \sup_{x,y\in\mathbb{R}^3}\frac{|\rho^a(x)u(x) - \rho^a(y)u(y)|}{|x-y|^{\mu}}.$$

Then the space of all functions for which

$$\partial^{\alpha}u \in C^{\mu}_{a+|\alpha|}(\mathbb{R}^3),\ \ |\alpha| \le l,$$

where l is a nonnegative integer is being denoted as $C_a^{\mu+l}(\mathbb{R}^3)$. Let $P(s)$ be the set of polynomials of three variables of the order less or equal to s for $s \ge 0$ and $P(s)$ is empty when $s < 0$. We make the following assumption on the right side of the linearized Cahn-Hilliard problem.

Assumption 3. *Let $f(x) \in C^{\mu}_{6+\varepsilon}(\mathbb{R}^3)$ for some $0 < \varepsilon < 1$ and the orthogonality relation*

$$(f(x), p(x))_{L^2(\mathbb{R}^3)} = 0 \tag{8}$$

holds for any polynomial $p(x) \in P(3)$ satisfying the equation $\Delta p(x) = 0$.

Remark. *A good example of such polynomials of the third order is*

$$\frac{a}{2}x_1^3 + \frac{b}{2}x_1x_2^2 + \frac{c}{2}x_1x_3^2,$$

where a, b and c are constants, such that $3a + b + c = 0$. The set of admissible $p(x)$ includes also constants, linear functions of three variables and many more examples.

By means of Lemma 2.3 of [17], under our Assumption 1 above on the potential function, the operator $-\Delta - V(x) - a$ is self-adjoint and unitarily equivalent to $-\Delta - a$ on $L^2(\mathbb{R}^3)$ via the wave operators (see [10, 14])

$$\Omega^{\pm} := s - \lim_{t\to\mp\infty}e^{it(-\Delta-V)}e^{it\Delta}$$

with the limit understood in the strong L^2 sense (see, e.g., [13] p.34, [4] p.90). Therefore, $-\Delta - V(x) - a$ on $L^2(\mathbb{R}^3)$ has only the essential spectrum $\sigma_{ess}(-\Delta - V(x) - a) = [-a,\ \infty)$. Via the spectral theorem, its functions of the continuous spectrum satisfying

$$[-\Delta - V(x)]\varphi_k(x) = k^2\varphi_k(x), \quad k \in \mathbb{R}^3, \tag{9}$$

in the integral formulation the Lippmann-Schwinger equation for the perturbed plane waves (see, e.g., [13] p.98)

$$\varphi_k(x) = \frac{e^{ikx}}{(2\pi)^{\frac{3}{2}}} + \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{e^{i|k||x-y|}}{|x-y|}(V\varphi_k)(y)dy \tag{10}$$

and the orthogonality relations

$$(\varphi_k(x), \varphi_q(x))_{L^2(\mathbb{R}^3)} = \delta(k-q), \quad k, q \in \mathbb{R}^3 \tag{11}$$

form the complete system in $L^2(\mathbb{R}^3)$. We introduce the following auxiliary functional space (see also [19, 20])

$$\tilde{W}^{2,\infty}(\mathbb{R}^3) := \{w(x) : \mathbb{R}^3 \to \mathbb{C} \mid w, \nabla w, \Delta w \in L^\infty(\mathbb{R}^3)\}. \tag{12}$$

As distinct from the standard Sobolev space we require here not the boundedness of all second partial derivatives of the function but of its Laplacian. Our main result is as follows.

Theorem 4. *Let Assumptions 1, 2 and 3 hold, $a \geq 0$ and $v_0(x)$ is given by (5). Then problem (3) admits a unique solution $u_a \in H^4(\mathbb{R}^3)$ if and only if*

$$(v_0(x), w(x))_{L^2(\mathbb{R}^3)} = 0 \tag{13}$$

for any $w(x) \in \tilde{W}^{2,\infty}(\mathbb{R}^3)$ satisfying equation (6).

Remark. *Note that $\varphi_k(x) \in \tilde{W}^{2,\infty}(\mathbb{R}^3)$, $k \in \mathbb{R}^3$, which was proven in Lemma A3 of [19]. By means of (9) these perturbed plane waves satisfy the homogeneous problem (6) when the wave vector k belongs to the sphere in three dimensions centered at the origin of radius $\sqrt{a}$.*

2. Proof of the Main Result

Armed with the technical lemma of the Appendix we proceed to prove the main result.

Proof of Theorem 4. The linearized Cahn-Hillard equation (3) is equivalent to system (4) in which the first equation admits a solution $v_0(x)$ given by (5). The function $v_0(x) \in L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ and $|x|v_0(x) \in L^1(\mathbb{R}^3)$ by means of Lemma A1 and Assumption 3. Then according to Theorem 3 of [20] the second equation in system (4) with $v_0(x)$ in its right side admits a unique solution in $H^2(\mathbb{R}^3)$ if and only if the orthogonality relation (13) holds. This solution of problem (3) $u_a(x) \in H^2(\mathbb{R}^3) \subset L^\infty(\mathbb{R}^3)$ via the Sobolev embedding theorem, $a \geq 0$ satisfies the equation

$$-\Delta u_a - V(x)u_a - au_a = v_0(x).$$

We use the formula

$$\Delta(Vu_a) = V\Delta u_a + 2\nabla V.\nabla u_a + u_a\Delta V \tag{14}$$

with the "dot" denoting the standard scalar product of two vectors in three dimensions. The first term in the right side of (14) is square integrable since $V(x)$ is bounded and $\Delta u_a(x) \in L^2(\mathbb{R}^3)$. Similarly $u_a\Delta V \in L^2(\mathbb{R}^3)$ since $u_a(x)$ is bounded and ΔV is square integrable by means of Assumption 2. For the second term in the right side of (14) we have $\nabla u_a(x) \in L^2(\mathbb{R}^3)$ and ∇V is bounded via Assumption 2, which yields $\nabla V.\nabla u_a \in L^2(\mathbb{R}^3)$ and therefore, $\Delta(Vu_a) \in L^2(\mathbb{R}^3)$. The right side of problem (3) belongs to $L^2(\mathbb{R}^3)$ due to Assumption 3. Indeed, since $\sup_{x\in\mathbb{R}^3}|\rho^{6+\varepsilon} f| \le C$, we arrive at the estimate

$$|f(x)| \le \frac{C}{(\rho(x))^{6+\varepsilon}},\ x \in \mathbb{R}^3 \tag{15}$$

with $\rho(x)$ defined explicitly in (7). Hence from equation (3) we deduce that $\Delta^2 u_a \in L^2(\mathbb{R}^3)$. Any partial third derivative of u_a is also square integrable due to the trivial estimate in terms of the $L^2(\mathbb{R}^3)$ norms of u_a and $\Delta^2 u_a$, which are finite. This implies that $u_a \in H^4(\mathbb{R}^3)$.

To investigate the issue of uniqueness we suppose $u_1, u_2 \in H^4(\mathbb{R}^3)$ are two solutions of problem (3). Then their difference $u(x) = u_1(x) - u_2(x) \in H^4(\mathbb{R}^3)$ satisfies equation (3) with vanishing right side. Clearly $u, \Delta u \in L^2(\mathbb{R}^3)$ and $Vu \in L^2(\mathbb{R}^3)$. Therefore, $v(x) = -\Delta u - V(x)u - au \in L^2(\mathbb{R}^3)$ and solves the equation $\Delta v = 0$. Since the Laplace operator does not have any nontrivial square integrable zero modes, $v(x) = 0$ a.e. in $\mathbb{R}^3$. Hence, we arrive at the homogeneous problem $(-\Delta - V(x) - a)u = 0, \quad u(x) \in L^2(\mathbb{R}^3)$. The operator in brackets is unitarily equivalent to $-\Delta - a$ on $L^2(\mathbb{R}^3)$ as discussed above and therefore $u(x) = 0$ a.e. in $\mathbb{R}^3$. □

3. Appendix

LEMMA A1. *Let Assumption 3 hold. Then* $v_0(x) \in L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ *and* $xv_0(x) \in L^1(\mathbb{R}^3)$.

Proof. According to the result of [2], for the solution of the Poisson equation (5) under the condition $f(x) \in C^{\mu}_{6+\varepsilon}(\mathbb{R}^3)$ and orthogonality relation (8) given in Assumption 3 we have $v_0(x) \in C^{\mu+2}_{4+\varepsilon}(\mathbb{R}^3)$. Hence $\sup_{x\in\mathbb{R}^3}|\rho^{4+\varepsilon} v_0| \le C$, such that

$$|v_0(x)| \le \frac{C}{(\rho(x))^{4+\varepsilon}},\ x \in \mathbb{R}^3.$$

The statement of the lemma easily follows from definition (7). □

Remark. *Note that the boundedness of $v_0(x)$ can be easily shown via the argument of Lemma 2.1 of [17], which relies on Young's inequality. The square integrability of $v_0(x)$ can be proven by applying the Fourier transform to it, using the facts that $f(x) \in L^2(\mathbb{R}^3)$, $|x|f(x) \in L^1(\mathbb{R}^3)$ and its Fourier image vanishes at the origin since it is orthogonal to a constant by means of Assumption 3.*

References

[1] N.D. Alikakos and G. Fusco, *Slow dynamics for the Cahn-Hilliard equation in higher space dimensions. I. Spectral estimates.*, Comm. Part. Diff. Eq. **19** (1994), 1397–1447.

[2] N. Benkirane, *Propriété d'indice en théorie Holderienne pour des opérateurs elliptiques dans $\mathbb{R}^n$*, CRAS **307** (1988), 577–580.

[3] L.A. Caffarelli and N.E. Muler, *An L^∞ bound for solutions of the Cahn-Hilliard equation*, Arch. Rational Mech. Anal. **133** (1995), 129–144.

[4] H.L. Cycon, R.G. Froese, W. Kirsch and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*, Springer, Berlin (1987).

[5] A. Ducrot, M. Marion and V. Volpert, *Systemes de réaction-diffusion sans propriété de Fredholm*, CRAS **340** (2005), 659–664.

[6] P. Howard, *Spectral analysis of stationary solutions of the Cahn-Hilliard equation*, Adv. Diff. Eq. **14** (2009), 87–120.

[7] A. Ducrot, M. Marion and V. Volpert, *Reaction-diffusion problems with non Fredholm operators*, Adv. Diff. Eq. **13** (2008), 1151–1192.

[8] P.J. Flory, *Thermodynamics of high polymer solutions*, J. Chem. Phys. **10** (1942), 51–61.

[9] B.L.G. Jonsson, M. Merkli, I.M. Sigal and F. Ting, *Applied analysis*, in preparation.

[10] T. Kato, *Wave operators and similarity for some non-selfadjoint operators*, Math. Ann. **162** (1966), 258–279.

[11] M.D. Korzec, P.L. Evans, A. Münch and B. Wagner, *Stationary solutions of driven fourth- and sixth-order Cahn-Hilliard-type equations*, SIAM J. Appl. Math. **69** (2008), 348–374.

[12] E. Lieb and M. Loss, *Analysis*, Graduate studies in mathematics series, volume 14, American Mathematical Society, Providence (1997).

[13] M. Reed and B. Simon, *Methods of modern mathematical physics. III: Scattering theory*, Academic Press, New York (1979).

[14] I. Rodnianski and W. Schlag, *Time decay for solutions of Schrödinger equations with rough and time-dependent potentials*, Invent. Math. **155** (2004), 451–513.

[15] V. Volpert, *Elliptic partial differential equations. Vol. 1: Fredholm property of elliptic problems in unbounded domains*, Birkhäuser, Berlin (2011).

[16] V. Volpert, B. Kazmierczak, M. Massot and Z.Peradzynski, *Solvability conditions for elliptic problems with non-Fredholm operators*, Appl. Math. **29** (2002), 219–238.

[17] V. Vougalter and V. Volpert, *Solvability conditions for some non-Fredholm operators*, Proc. Edinb. Math. Soc. **54** (2011), 249–271.
[18] V. Vougalter and V. Volpert, *On the solvability conditions for some non Fredholm operators*, Int. J. Pure Appl. Math., **60** (2010), 169–191.
[19] V. Vougalter and V. Volpert, *On the solvability conditions for the diffusion equation with convection terms*, Commun. Pure Appl. Anal. **11** (2012), 365–373.
[20] V. Vougalter and V. Volpert, *Solvability relations for some non Fredholm operators*, Int. Electron. J. Pure Appl. Math., **2** (2010), 75–83.

Authors' addresses:

Vitaly Volpert
Institute Camille Jordan, UMR 5208 CNRS
University Lyon 1
Villeurbanne, 69622, France
E-mail: volpert@math.univ-lyon1.fr

Vitali Vougalter
Department of Mathematics and Applied Mathematics
University of Cape Town
Private Bag, Rondebosch 7701, South Africa
E-mail: Vitali.Vougalter@uct.ac.za

Received October 19, 2010
Revised February 24, 2011

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 11–30

Non-vanishing Theorems for Rank Two Vector Bundles on Threefolds[1]

Edoardo Ballico, Paolo Valabrega and Mario Valenzano

Abstract. *The paper investigates the non-vanishing of $H^1(\mathcal{E}(n))$, where $\mathcal{E}$ is a (normalized) rank two vector bundle over any smooth irreducible threefold X with $\mathrm{Pic}(X) \cong \mathbb{Z}$. If ϵ is defined by the equality $\omega_X = \mathcal{O}_X(\epsilon)$, and α is the least integer t such that $H^0(\mathcal{E}(t)) \neq 0$, then, for a non-stable $\mathcal{E}$, $H^1(\mathcal{E}(n))$ does not vanish at least between $\frac{\epsilon - c_1}{2}$ and $-\alpha - c_1 - 1$. The paper also shows that there are other non-vanishing intervals, whose endpoints depend on α and on the second Chern class of $\mathcal{E}$. If $\mathcal{E}$ is stable $H^1(\mathcal{E}(n))$ does not vanish at least between $\frac{\epsilon - c_1}{2}$ and $\alpha - 2$. The paper considers also the case of a threefold X with $\mathrm{Pic}(X) \neq \mathbb{Z}$ but $\mathrm{Num}(X) \cong \mathbb{Z}$ and gives similar non-vanishing results.*

Keywords: Rank Two Vector Bundles, Smooth Threefolds, Non-vanishing of 1-Cohomology.
MS Classification 2010: 14J60, 14F05

1. Introduction

In 1942 G. Gherardelli ([5]) proved that, if C is a smooth irreducible curve in $\mathbb{P}^3$ whose canonical divisors are cut out by the surfaces of some degree e and moreover all linear series cut out by the surfaces in $\mathbb{P}^3$ are complete, then C is the complete intersection of two surfaces. Shortly and in the language of modern algebraic geometry: every e-subcanonical smooth curve C in $\mathbb{P}^3$ such that $h^1(\mathcal{I}_C(n)) = 0$ for all n is the complete intersection of two surfaces.

Thanks to the Serre correspondence between curves and vector bundles (see [7, 8, 9]) the above statement is equivalent to the following one: if $\mathcal{E}$ is a rank two vector bundle on $\mathbb{P}^3$ such that $h^1(\mathcal{E}(n)) = 0$ for all n, then $\mathcal{E}$ splits.

---

[1]The paper was written while all authors were members of INdAM-GNSAGA.
Lavoro eseguito con il supporto del progetto PRIN “Geometria delle varietà algebriche e dei loro spazi di moduli”, cofinanziato dal MIUR (cofin 2008).

There are many improvements of the above result with a variety of different approaches (see for instance [2, 3, 4, 13, 15]): it comes out that a rank two vector bundle $\mathcal{E}$ on $\mathbb{P}^3$ is forced to split if $h^1(\mathcal{E}(n))$ vanishes for just one strategic n, and such a value n can be chosen arbitrarily within a suitable interval, whose endpoints depend on the Chern classes and the least number α such that $h^0(\mathcal{E}(\alpha)) \neq 0$.

When rank two vector bundles on a smooth threefold X of degree d in $\mathbb{P}^4$ are concerned, similar results can be obtained, with some interesting difference.

In 1998 Madonna ([11]) proved that on a smooth threefold X of degree d in $\mathbb{P}^4$ there are ACM rank two vector bundles (i.e. whose 1-cohomology vanishes for all twists) that do not split. And this can happen, for a normalized vector bundle $\mathcal{E}$ ($c_1 \in \{0, -1\}$), only when $1 - \frac{d+c_1}{2} < \alpha < \frac{d-c_1}{2}$, while an ACM rank two vector bundle on X whose α lies outside of the interval is forced to split.

The following non-vanishing results for a normalized non-split rank two vector bundle on a smooth irreducible thereefold of degree d in $\mathbb{P}^4$ are proved in [11]:

- if $\alpha \leq 1 - \frac{d+c_1}{2}$, then $h^1(\mathcal{E}(\frac{d-3-c_1}{2})) \neq 0$ if $d + c_1$ is odd, $h^1(\mathcal{E}(\frac{d-4-c_1}{2})) \neq 0, h^1(\mathcal{E}(\frac{d-2-c_1}{2})) \neq 0$ if $d + c_1$ is even, while $h^1(\mathcal{E}(\frac{d-c_1}{2})) \neq 0$ if $d + c_1$ is even and moreover $\alpha \leq -\frac{d+c_1}{2}$;
- if $\alpha \geq \frac{d-c_1}{2}$, then $h^1(\mathcal{E}(\frac{d-3-c_1}{2})) \neq 0$ if $d + c_1$ is odd, while $h^1(\mathcal{E}(\frac{d-4-c_1}{2})) \neq 0$ if $d + c_1$ is even.

In [11] it is also claimed that the same techniques work to obtain similar non-vanishing results on any smooth threefold X with $\mathrm{Pic}(X) \cong \mathbb{Z}$ and $h^1(\mathcal{O}_X(n)) = 0$, for every n.

The present paper investigates the non-vanishing of $H^1(\mathcal{E}(n))$, where $\mathcal{E}$ is a rank two vector bundle over any smooth irreducible threefold X such that $\mathrm{Pic}(X) \cong \mathbb{Z}$ and $H^1(\mathcal{O}_X(n)) = 0$, for all n. Actually we can prove that for such an $\mathcal{E}$ there is a wider range of non-vanishing for $h^1(\mathcal{E}(n))$, so improving the above results.

More precisely, when $\mathcal{E}$ is (normalized and) non-stable ($\alpha \leq 0$) the first cohomology module does not vanish at least between the endpoints $\frac{\epsilon - c_1}{2}$ and $-\alpha - c_1 - 1$, where ϵ is defined by the equality $\omega(X) = \mathcal{O}_X(\epsilon)$ (and is $d - 5$ if $X \subset \mathbb{P}^4$, where $d = \deg(X)$). But we can show that there are other non-vanishing intervals, whose endpoints depend on α and also on the second Chern class c_2 of $\mathcal{E}$.

If on the contrary $\mathcal{E}$ is stable the first cohomology module does not vanish at least between the endpoints $\frac{\epsilon - c_1}{2}$ and $\alpha - 2$, but other ranges of non-vanishing can be produced.

We give a few examples obtained by pull-back from vector bundles on $\mathbb{P}^3$.

We must remark that most of our non-vanishing results do not exclude the range for α between the endpoints $1 - \frac{d+c_1}{2}$ and $\frac{d-c_1}{2}$ (for a general threefold it becomes $-\frac{\epsilon+3+c_1}{2} < \alpha < \frac{\epsilon+5-c_1}{2}$). Actually [11] produces some examples of non-split ACM rank two vector bundles on smooth hypersurfaces in $\mathbb{P}^4$, but it can be seen that they do not conflict with our theorems.

As to threefolds with $\mathrm{Pic}(X) \neq \mathbb{Z}$, we need to observe that a key point is a good definition of the integer α. We are able to prove, by using a boundedness argument, that α exists when $\mathrm{Pic}(X) \neq \mathbb{Z}$ but $\mathrm{Num}(X) \cong \mathbb{Z}$. In this event the correspondence between rank two vector bundles and two-codimensional subschemes can be proved to hold. In order to obtain non-vanishing results that are similar to the results proved when $\mathrm{Pic}(X) \cong \mathbb{Z}$, we need also use the Kodaira vanishing theorem, which holds in characteristic 0. We can extend the results to characteristic $p > 0$ if we assume a Kodaira-type vanishing condition.

In this paper we investigate non-vanishing theorems for rank two vector bundles on *any* threefold. The problem looks quite different if the threefold is general of belongs to some family (for the case of ACM bundles see for instance [14] and [1]).

Moreover we observe that our examples of section 6 are sharp but the threefolds (except one) are quadric hypersurfaces, so that one can guess that some stronger statement holds when the degree d is large enough.

2. Notation

We work over an algebraically closed field $\mathbf{k}$ of any characteristic.

Let X be a non-singular irreducible projective algebraic variety of dimension 3, for short a smooth threefold. We fix an ample divisor H on X, so we consider the polarized threefold (X, H). We denote with $\mathcal{O}_X(n)$, instead of $\mathcal{O}_X(nH)$, the invertible sheaf corresponding to the divisor nH, for each $n \in \mathbb{Z}$.

For every cycle Z on X of codimension i it is defined its degree with respect to H, i.e. $\deg(Z; H) := Z \cdot H^{3-i}$, having identified a codimension 3 cycle on X, i.e. a 0-dimensional cycle, with its degree, which is an integer.

From now on (with the exception of section 7) we consider a smooth polarized threefold $(X, \mathcal{O}_X(1)) = (X, H)$ that satifies the following conditions:

(**C1**) $\mathrm{Pic}(X) \cong \mathbb{Z}$ generated by $[H]$,

(**C2**) $H^1(X, \mathcal{O}_X(n)) = 0$ for every $n \in \mathbb{Z}$,

(**C3**) $H^0(X, \mathcal{O}_X(1)) \neq 0$.

By condition (**C1**) every divisor on X is linearly equivalent to aH for some integer $a \in \mathbb{Z}$, i.e. every invertible sheaf on X is (up to an isomorphism) of type $\mathcal{O}_X(a)$ for some $a \in \mathbb{Z}$, in particular we have for the canonical divisor $K_X \sim \epsilon H$, or equivalently $\omega_X \simeq \mathcal{O}_X(\epsilon)$, for a suitable integer ϵ. Furthermore, by Serre duality condition (**C2**) implies that $H^2(X, \mathcal{O}_X(n)) = 0$ for all $n \in \mathbb{Z}$.

Since by assumption $A^1(X) = \mathrm{Pic}(X)$ is isomorphic to $\mathbb{Z}$ through the map $[H] \mapsto 1$, where $[H] = c_1(\mathcal{O}_X(1))$, we identify the first Chern class $c_1(\mathcal{F})$ of a coherent sheaf with a whole number c_1, where $c_1(\mathcal{F}) = c_1 H$.

The second Chern class $c_2(\mathcal{F})$ gives the integer $c_2 = c_2(\mathcal{F}) \cdot H$ and we will call this integer the second Chern number or the second Chern class of $\mathcal{F}$.

We set

$$d := \deg(X; H) = H^3,$$

so d is the "degree" of the threefold X with respect to the ample divisor H.

Let $c_1(X)$ and $c_2(X)$ be the first and second Chern classes of X, that is of its tangent bundle TX (which is a locally free sheaf of rank 3); then we have

$$c_1(X) = [-K_X] = -\epsilon[H],$$

so we identify the first Chern class of X with the integer $-\epsilon$. Moreover we set

$$\tau := \deg(c_2(X); H) = c_2(X) \cdot H,$$

i.e. τ is the degree of the second Chern class of the threefold X.

In the following we will call the triple of integers (d, ϵ, τ) the *characteristic numbers* of the polarized threefold $(X, \mathcal{O}_X(1))$.

We recall the well-known Riemann-Roch formula on the threefold X (e.g. see [18], Proposition 4).

Theorem 2.1 (Riemann-Roch). *Let $\mathcal{F}$ be a rank r coherent sheaf on X with Chern classes $c_1(\mathcal{F})$, $c_2(\mathcal{F})$ and $c_3(\mathcal{F})$. Then the Euler-Poincaré characteristic of $\mathcal{F}$ is*

$$\begin{aligned}\chi(\mathcal{F}) =& \frac{1}{6}\Big(c_1(\mathcal{F})^3 - 3c_1(\mathcal{F}) \cdot c_2(\mathcal{F}) + 3c_3(\mathcal{F})\Big) + \frac{1}{4}\Big(c_1(\mathcal{F})^2 - 2c_2(\mathcal{F})\Big) \cdot c_1(X) \\ &+ \frac{1}{12} c_1(\mathcal{F}) \cdot \Big(c_1(X)^2 + c_2(X)\Big) + \frac{r}{24} c_1(X) \cdot c_2(X)\end{aligned}$$

where $c_1(X)$ and $c_2(X)$ are the Chern classes of X, that is the Chern classes of the tangent bundle TX of X.

So applying the Riemann-Roch Theorem to the invertible sheaf $\mathcal{O}_X(n)$, for each $n \in \mathbb{Z}$, we get the Hilbert polynomial of the sheaf $\mathcal{O}_X(1)$

$$\chi(\mathcal{O}_X(n)) = \frac{d}{6}\left(n - \frac{\epsilon}{2}\right)\left[\left(n - \frac{\epsilon}{2}\right)^2 + \frac{\tau}{2d} - \frac{\epsilon^2}{4}\right]. \tag{1}$$

Let $\mathcal{E}$ be a rank 2 vector bundle on the threefold X with Chern classes $c_1(\mathcal{E})$ and $c_2(\mathcal{E})$, i.e. with Chern numbers c_1 and c_2. We assume that $\mathcal{E}$ is normalized, i.e. that $c_1 \in \{0, -1\}$. It is defined the integer α, the so called *first relevant* *level*, such that $h^0(\mathcal{E}(\alpha)) \neq 0, h^0(\mathcal{E}(\alpha-1)) = 0$. If $\alpha > 0$, $\mathcal{E}$ is called stable, non-stable otherwise. We set

$$\vartheta = \frac{3c_2}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} - \frac{3c_1^2}{4}, \qquad \zeta_0 = \frac{\epsilon - c_1}{2}, \quad \text{and} \quad w_0 = [\zeta_0] + 1,$$

where $[\zeta_0]$ = integer part of ζ_0, so the Hilbert polynomial of $\mathcal{E}$ can be written as

$$\chi(\mathcal{E}(n)) = \frac{d}{3}\big(n - \zeta_0\big)\Big[\big(n-\zeta_0\big)^2 - \vartheta\Big]. \tag{2}$$

If $\vartheta \geq 0$ we set

$$\zeta = \zeta_0 + \sqrt{\vartheta}$$

so in this case the Hilbert polynomial of $\mathcal{E}$ has the three real roots $\zeta' \leq \zeta_0 \leq \zeta$ where $\zeta' = \zeta_0 - \sqrt{\vartheta}$. We also define $\bar{\alpha} = [\zeta] + 1$.

The polinomial $\chi(\mathcal{E}(n))$, as a rational polynomial, has three real roots if and only if $\vartheta \geq 0$, and it has only one real root if and only if $\vartheta < 0$.

If $\mathcal{E}$ is normalized, we set

$$\delta = c_2 + c_1 d\alpha + d\alpha^2.$$

Proposition 2.2. *It holds $\delta = 0$ if and only if $\mathcal{E}$ splits.*

Proof. (see also [17], Lemma 3.13) In fact, if $\mathcal{E} = \mathcal{O}_X(a) \otimes \mathcal{O}_X(-a + c_1)$, for some $a \geq 0$, then a direct computation shows that $\delta = 0$. Conversely, if $\mathcal{E}$ is a non-split bundle, then $\mathcal{E}(\alpha)$ has a non-vanishing section that gives rise to a two-codimensional scheme, whose degree, by [6], Appendix A, 3, C6, is δ, which cannot be 0. □

Unless stated otherwise, we work over the smooth polarized threefold X and $\mathcal{E}$ *is a normalized non-split rank two vector bundle on X.*

3. About the Characteristic Numbers ϵ and τ

In this section we want to recall some essentially known properties of the characteristic numbers of the threefold X (see also [16] for more general statements). We start with the following remark.

Remark 3.1. *For the fixed ample invertible sheaf $\mathcal{O}_X(1)$ we have:*

$$h^0(\mathcal{O}_X(n)) = 0 \text{ for } n < 0, \quad h^0(\mathcal{O}_X) = 1, \quad h^0(\mathcal{O}_X(n)) \neq 0 \text{ for } n > 0,$$

and also $h^0(\mathcal{O}_X(m)) - h^0(\mathcal{O}_X(n)) > 0$ for all $n, m \in \mathbb{Z}$ with $m > n \geq 0$.

Moreover it holds

$$\chi(\mathcal{O}_X) = h^0(\mathcal{O}_X) - h^3(\mathcal{O}_X) = 1 - h^0(\mathcal{O}_X(\epsilon)),$$

so we have:

$$\chi(\mathcal{O}_X) = 1 \iff \epsilon < 0, \quad \chi(\mathcal{O}_X) = 0 \iff \epsilon = 0, \quad \chi(\mathcal{O}_X) < 0 \iff \epsilon > 0.$$

Proposition 3.2. *Let $(X, \mathcal{O}_X(1))$ be a smooth polarized threefold with characteristic numbers (d, ϵ, τ). Then it holds:*

1) $\epsilon \geq -4$,
2) $\epsilon = -4$ *if and only if* $X = \mathbb{P}^3$, *i.e.* $(d, \epsilon, \tau) = (1, -4, 6)$ *and so* $\frac{\tau}{2d} - \frac{\epsilon^2}{4} = -1$,
3) *if* $\epsilon = -3$, *then* $(d, \epsilon, \tau) = (2, -3, 8)$ *and* $\frac{\tau}{2d} - \frac{\epsilon^2}{4} = -\frac{1}{4}$,
4) $\epsilon\tau$ *is a multiple of* 24, *in particular if* $\epsilon < 0$ *then* $\epsilon\tau = -24$ *and moreover the only possibilities for* (ϵ, τ) *are the following:*
$$(\epsilon, \tau) \in \{(-4, 6), (-3, 8), (-2, 12), (-1, 24)\},$$
5) *if* $\epsilon \neq 0$, *then* $\tau > 0$,
6) *if* $\epsilon = 0$, *then* $\tau > -2d$,
7) τ *is always even,*
8) *if* ϵ *is even, then* $\frac{\tau}{2d} - \frac{\epsilon^2}{4} \geq -1$,
9) *if* ϵ *is odd, then* $\frac{\tau}{2d} - \frac{\epsilon^2}{4} \geq -\frac{1}{4}$.

Proof. For statements 1), 2), 3) see [16].

4) Observe that $\chi(\mathcal{O}_X) = -\frac{1}{24}\epsilon\tau$ is an integer, and moreover, if $\epsilon < 0$, then $\chi(\mathcal{O}_X) = 1$. If $\epsilon < 0$, then by 1) we have $\epsilon \in \{-4, -3, -2, -1\}$ and so we obtain the thesis.
5) By Remark 3.1 we have: if $\epsilon > 0$ then $-\frac{1}{24}\epsilon\tau < 0$, while if $\epsilon < 0$ then $-\frac{1}{24}\epsilon\tau > 0$. In both cases we deduce $\tau > 0$.
6) If $\epsilon = 0$, then we have
$$\chi(\mathcal{O}_X(n)) = \frac{d}{6}n\left(n^2 + \frac{\tau}{2d}\right),$$
and also
$$\chi(\mathcal{O}_X(n)) = h^0(\mathcal{O}_X(n)) > 0 \quad \forall n > 0,$$
therefore we must have $\frac{2d+\tau}{12} > 0$, so $\tau > -2d$.
7) Assume that ϵ is even, then we have
$$d\left(1 - \frac{\epsilon}{2}\right)\left(1 + \frac{\epsilon}{2}\right) + \frac{\tau}{2} = d\left(1 - \frac{\epsilon^2}{4} + \frac{\tau}{2d}\right) = 6\,\chi\left(\mathcal{O}_X\left(\frac{\epsilon}{2} + 1\right)\right) \in \mathbb{Z}$$
and moreover $d\left(1 - \frac{\epsilon}{2}\right)\left(1 + \frac{\epsilon}{2}\right) \in \mathbb{Z}$, so τ must be even.
If ϵ is odd, the proof is quite similar.

8) Let ϵ be even. If it holds

$$h^0\left(\mathcal{O}_X\left(\frac{\epsilon}{2}+1\right)\right)-h^0\left(\mathcal{O}_X\left(\frac{\epsilon}{2}-1\right)\right)=\chi\left(\mathcal{O}_X\left(\frac{\epsilon}{2}+1\right)\right)<0,$$

then we must have $h^0\left(\mathcal{O}_X\left(\frac{\epsilon}{2}-1\right)\right)\neq 0$, which implies

$$h^0\left(\mathcal{O}_X\left(\frac{\epsilon}{2}+1\right)\right)-h^0\left(\mathcal{O}_X\left(\frac{\epsilon}{2}-1\right)\right)\geq 0,$$

a contradiction. So we must have

$$\chi\left(\mathcal{O}_X\left(\frac{\epsilon}{2}+1\right)\right)=\frac{d}{6}\left(1+\frac{\tau}{2d}-\frac{\epsilon^2}{4}\right)\geq 0,$$

therefore

$$\frac{\tau}{2d}-\frac{\epsilon^2}{4}\geq -1.$$

9) The proof is quite similar to the proof of 8). □

4. Non-stable Vector Bundles ($\alpha \leq 0$)

We make the following assumption:

$\mathcal{E}$ is a normalized non-split rank two vector bundle with $\alpha \leq 0$.

LEMMA 4.1. *For every integer n it holds:*

$$\chi(\mathcal{O}_X(n-\alpha))-\chi(\mathcal{O}_X(\epsilon-n-\alpha-c_1))-\chi(\mathcal{E}(n))=(n-\zeta_0)\delta.$$

Proof. It is a straightforward computation using formulas (1) and (2) for the Hilbert polynomial of $\mathcal{O}_X(1)$ and $\mathcal{E}$, respectively. □

PROPOSITION 4.2. *Assume that $\zeta_0 < -\alpha - c_1 - 1$. Then it holds:*

$$h^1(\mathcal{E}(n))-h^2(\mathcal{E}(n))=(n-\zeta_0)\delta$$

for every integer n such that $\zeta_0 < n \leq -\alpha - c_1 - 1$.

Proof. For each n such that $\zeta_0 < n \leq -\alpha - c_1 - 1$ it holds: $\epsilon - n + \alpha < -1$ and $n + \alpha + c_1 \leq -1$, so we have

$$h^3(\mathcal{O}_X(n-\alpha)) = h^0(\mathcal{O}_X(\epsilon-n+\alpha)) = 0$$
$$h^3(\mathcal{O}_X(\epsilon-n-\alpha-c_1)) = h^0(\mathcal{O}_X(n+\alpha+c_1)) = 0,$$

therefore we obtain:

$$\begin{aligned}
h^0(\mathcal{E}(n)) &= h^0(\mathcal{O}_X(n-\alpha)) = \chi(\mathcal{O}_X(n-\alpha)) \\
h^3(\mathcal{E}(n)) &= h^0(\mathcal{E}(\epsilon - n - c_1)) = h^0(\mathcal{O}_X(\epsilon - n - \alpha - c_1)) \\
&\qquad\qquad\qquad\qquad\quad = \chi(\mathcal{O}_X(\epsilon - n - \alpha - c_1)).
\end{aligned}$$

Hence

$$\begin{aligned}
h^1(\mathcal{E}(n)) - h^2(\mathcal{E}(n)) &= h^0(\mathcal{E}(n)) - h^3(\mathcal{E}(n)) - \chi(\mathcal{E}(n)) = \\
&= \chi(\mathcal{O}_X(n-\alpha)) - \chi(\mathcal{O}_X(\epsilon - n - \alpha - c_1)) - \chi(\mathcal{E}(n)),
\end{aligned}$$

so using Lemma 4.1 we obtain tha claim. □

THEOREM 4.3. *Let us assume that $\zeta_0 < -\alpha - c_1 - 1$ and let n be such that $\zeta_0 < n \leq -\alpha - 1 - c_1$. Then $h^1(\mathcal{E}(n)) \geq (n - \zeta_0)\delta$. In particular $h^1(\mathcal{E}(n)) \neq 0$.*

Proof. It is enough to observe that $h^1(\mathcal{E}(n)) - h^2(\mathcal{E}(n)) = (n - \zeta_0)\delta$, by Proposition 4.2, and that the right side of this equality is strictly positive for a non-split vector bundle. □

REMARK 4.4. *Observe that the above theorem describes a non-empty set of integers if and only if $-\alpha - c_1 - 1 > \zeta_0$; this means $\alpha < -\frac{\epsilon+2+c_1}{2}$, i.e. $\alpha \leq -\frac{\epsilon+3+c_1}{2}$. So our assumption on α agrees with the bound of [11].*
Observe that the inequality on α implies that $\alpha \leq -2$ if $\epsilon \geq 1$.

The non-vanishing result above can be improved, if other invariants both of the threefold and the bundle are considered.

Now we set $\lambda = \frac{\tau}{2d} - \frac{\epsilon^2}{4}$ and consider the following degree 3 polynomial:

$$F(X) = X^3 + \left(\lambda - \frac{6\delta}{d}\right) X + \frac{6\delta}{d}\left(\alpha + \frac{c_1}{2}\right).$$

It is easy to see that, if $\frac{6\delta}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} \leq 0$, then $F(X)$ is strictly increasing and so it has only one real root X_0.

THEOREM 4.5. *Assume that $\frac{6\delta}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} \leq 0$. Let n be such that $\epsilon - \alpha - c_1 + 1 \leq n < -\alpha + X_0 + \zeta_0$, where X_0 = unique real root of $F(X)$. Then $h^1(\mathcal{E}(n)) \geq -\frac{d}{6} F\left(n + \alpha - \zeta_0 + \frac{c_1}{2}\right) > -\frac{d}{6} F(X_0) = 0$. In particular $h^1(\mathcal{E}(n)) \neq 0$.*

Proof. For each n such that $\epsilon - \alpha - c_1 + 1 \leq n < -\alpha + X_0 + \zeta_0$ it holds: $\epsilon - n + \alpha \leq -1$ and $\epsilon - n - c_1 \leq \alpha - 1$, so we have

$$\begin{aligned}
h^3(\mathcal{O}_X(n-\alpha)) &= h^0(\mathcal{O}_X(\epsilon - n + \alpha)) = 0 \\
h^3(\mathcal{E}(n)) &= h^0(\mathcal{E}(\epsilon - n - c_1)) = 0.
\end{aligned}$$

Moreover, taking into account the exact sequence

$$0 \to \mathcal{O}_X(n-\alpha) \to \mathcal{E}(n) \to \mathcal{I}_Z(n+\alpha) \to 0$$

which arises from the Serre correspondence (see [18], Theorem 4), and where Z is the zero-locus of a non-zero section of $\mathcal{E}(\alpha)$, we obtain:

$$h^0(\mathcal{E}(n)) \geq h^0(\mathcal{O}_X(n-\alpha)) = \chi(\mathcal{O}_X(n-\alpha)).$$

Hence

$$\begin{aligned} h^1(\mathcal{E}(n)) &= h^0(\mathcal{E}(n)) + h^2(\mathcal{E}(n)) - h^3(\mathcal{E}(n)) - \chi(\mathcal{E}(n)) \\ &\geq \chi(\mathcal{O}_X(n-\alpha)) - \chi(\mathcal{E}(n)) = \text{ (by Lemma 4.1)} \\ &= (n-\zeta_0)\delta + \chi(\mathcal{O}_X(\epsilon - n - \alpha - c_1)) \\ &= (n-\zeta_0)\delta - \frac{d}{6}\left(n+\alpha-\zeta_0+\frac{c_1}{2}\right)\left[\left(n+\alpha-\zeta_0+\frac{c_1}{2}\right)^2+\lambda\right], \end{aligned}$$

so, if we put $X = n+\alpha-\zeta_0+\frac{c_1}{2}$, then we obtain: $h^1(\mathcal{E}(n)) \geq -\frac{d}{6}F(X) > -\frac{d}{6}F(X_0) = 0$, because of the hypothesis $n < -\alpha + X_0 + \zeta_0$ and the fact that F is strictly increasing. □

The proofs of the above theorems work perfectly without any restriction on ϵ, while for the proof of the following theorem a few more words are required if $\epsilon \leq 0$.

Theorem 4.6. *Assume that* $\frac{6\delta}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} - \frac{3c_1^2}{4} \geq 0$. *Let* $n > \zeta_0$ *be such that* $\epsilon - \alpha - c_1 + 1 \leq n < \zeta_0 + \sqrt{\frac{6\delta}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} - \frac{3c_1^2}{4}}$ *and put*

$$S(n) = \frac{d}{6}\left(n - \frac{\epsilon - c_1}{2}\right)\left[\left(n - \frac{\epsilon - c_1}{2}\right)^2 - 6\frac{c_2 + d\alpha^2 + c_1 d\alpha}{d} + \frac{\tau}{2d} - \frac{\epsilon^2}{4} + \frac{3c_1^2}{4}\right].$$

Then $h^1(\mathcal{E}(n)) \geq -S(n) > 0$. *In particular* $h^1(\mathcal{E}(n)) \neq 0$.

Proof. **Case 1: $\boldsymbol{\epsilon \geq 1}$.** Assume $c_1 = 0$. Under our hypothesis $h^0(\mathcal{E}(\epsilon - n)) = 0$ and so $h^1(\mathcal{E}(n)) - h^2(\mathcal{E}(n)) \geq h^0(\mathcal{O}_X(n-\alpha)) - \chi(\mathcal{E}(n))$. Observe that $h^0(\mathcal{O}_X(n-\alpha)) - \chi(\mathcal{E}(n)) + S(n) = -\frac{1}{2}nd\alpha(-\epsilon+n+\alpha)) - \frac{1}{12}d\alpha(-3\epsilon\alpha + 2\alpha^2 + \epsilon^2 + \frac{\tau}{d}) \geq 0$ (by direct computation). Therefore we have: $h^1(\mathcal{E}(n)) \geq h^2(\mathcal{E}(n)) - S(n)$. Hence $h^1(\mathcal{E}(n))$ may possibly vanish when

$$\left(n - \frac{\epsilon}{2}\right)^2 - 6\frac{c_2 + d\alpha^2}{d} + \frac{\tau}{2d} - \frac{\epsilon^2}{4} \geq 0.$$

When $S(n) < 0$, so $-S(n) > 0$, $h^1(\mathcal{E}(n)) \geq -S(n) > 0$ and in particular it cannot vanish.

If $c_1 = -1$ the proof is quite similar.

Case 2: $\epsilon \leq 0$.

A. $\epsilon \leq -2$.
We need to know that

$$\frac{1}{2}nd\alpha(-\epsilon + n + \alpha) + \frac{1}{12}d\alpha(\epsilon^2 + \frac{\tau}{d} - 3\epsilon\alpha + 2\alpha^2) \leq 0.$$

The first term of the sum is for sure negative; as for

$$\frac{1}{12}d\alpha\left(\epsilon^2 + \frac{\tau}{d}\right) + \frac{1}{12}d\alpha^2(-3\epsilon + 2\alpha)$$

we observe that the quantity in brackets has discriminant

$$\Delta = \epsilon^2 - 8\frac{\tau}{d} = 4\left(\frac{\epsilon^2}{4} - \frac{\tau}{2d} + \frac{\tau}{2d} - 8\frac{\tau}{d}\right) \leq 4(1-15) < 0.$$

Therefore it is positive for all $\alpha \leq 0$ and the product is negative.

B. $\epsilon = -1$.
We need to know that

$$\frac{1}{2}nd\alpha(1 + n + \alpha) + \frac{1}{12}d\alpha\left(1 + \frac{\tau}{d}\right) + \frac{1}{12}d\alpha^2(3 + 2\alpha) \leq 0.$$

If $\alpha \leq -2$, then it is enough to observe that $\frac{\tau}{d} + 3\alpha + 2\alpha^2 \geq 0$. If $\alpha = -1$ we have to consider $-\frac{1}{2}n^2d + \frac{1}{12}d\frac{\tau}{d}$ and then we observe that $6n^2 + \frac{\tau}{d} > 0$. If $\alpha = 0$ obviously the quantity is 0.

C. $\epsilon = 0$.
In theorem 4.5 we need to know that

$$\frac{1}{2}nd\alpha(n + \alpha) + \frac{1}{12}d\alpha\left(\frac{\tau}{d}\right) + \frac{1}{12}d\alpha^2(2\alpha) \leq 0.$$

It is enough to observe that $n + \alpha \geq 1$ and that $2\alpha^2 + \frac{\tau}{d} > 0$ (by Proposition 3.2(6)), if $\alpha < 0$; otherwise we have a 0 quantity. □

Remark 4.7. *Observe that in Theorems 4.5 and 4.6 α can be zero.*

Remark 4.8. *Observe that the case $\alpha = 0$ in Theorem 4.3 can occur only if $\epsilon \leq -c_1 - 3$.*

Remark 4.9. *In theorem 4.6 we do not use the hypothesis* $-\frac{\epsilon+3}{2} \geq \alpha$*, but we assume that* $6\frac{c_2+d\alpha^2}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} - 1 \geq 0$*. In theorem 4.5 we do not use the hypothesis* $-\frac{\epsilon+3}{2} \geq \alpha$*, but we assume that* $6\frac{c_2+d\alpha^2}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} < 0$*. Moreover in both theorems there is a range for* n*, the left endpoint being* $\epsilon - \alpha - c_1 + 1$ *and the right endpoint being either* $\zeta_0 + \sqrt{6\frac{c_2+d\alpha^2}{d} - \frac{\tau}{2d} + \frac{\epsilon^2}{4} - 1}$ *(4.6) or* $\zeta_0 - \alpha + X_0$ *(4.5).*

In [11] there are examples of ACM non-split vector bundles on smooth threefolds in $\mathbb{P}^4$*, with* $-\frac{\epsilon+3+c_1}{2} < \alpha < \frac{\epsilon+5-c_1}{2}$*. We want to emphasize that our theorems do not conflict with the examples of [11]: if* C *is any curve described in [11] and lying on a smooth threefold of degree* d*, then our numerical constraints cannot be satisfied (we have checked it directly in many but not all cases).*

Remark 4.10. *Let us consider a smooth degree* d *threefold* $X \subset \mathbb{P}^4$*.*

We have:

$$\epsilon = d - 5, \quad \tau = d(10 - 5d + d^2), \quad \vartheta = \frac{3c_2}{d} - \frac{d^2 - 5 + 3c_1^2}{4}$$

(see [18]). As to the characteristic function of $\mathcal{O}_X$ *and* $\mathcal{E}$*, it holds:*

$$\chi(\mathcal{O}_X(n)) = \frac{d}{6}\left(n - \frac{d-5}{2}\right)\left[\left(n - \frac{d-5}{2}\right)^2 + \frac{d^2-5}{4}\right],$$

$$\chi(\mathcal{E}(n)) = \frac{d}{3}\left(n - \frac{d-5-c_1}{2}\right)\left[\left(n - \frac{d-5-c_1}{2}\right)^2 + \frac{d^2}{4} - \frac{5}{4} + \frac{3c_1^2}{4} - \frac{3c_2}{d}\right].$$

Then it is easy to see that the hypothesis of Theorem 4.6, i.e. $6\frac{\delta}{d} - \frac{d^2-5+3c_1^2}{4} \geq 0$ *is for sure fulfilled if* $c_2 \geq 0, \alpha \leq -\frac{d-2+c_1}{2}$*. In fact we have (for the sake of simplicity when* $c_1 = 0$*):* $-6\frac{6c_2+d\alpha^2}{d} + \frac{d^2-5}{4} \leq \frac{d^2-5}{4} - 6\frac{d^2-2d+1}{4} = -\frac{5d^2-12d+11}{4} < 0$*.*

Remark 4.11. *Condition* (**C2**) *holds for sure if* X *is a smooth hypersurface of* $\mathbb{P}^4$*. In general, for a characteristic* 0 *base field, only the Kodaira vanishing holds ([6], remark 7.15) and so, unless we work over a threefold* X *having some stronger vanishing, we need assume, in Theorems 4.3, 4.5, 4.6 that* $n - \alpha \notin \{0, \dots, \epsilon\}$ *(which implies, by duality, that also* $\epsilon - n + \alpha \notin \{0, \dots, \epsilon\}$*).*

Observe that the first assumption $(n - \alpha \notin \{0, \dots, \epsilon\})$ *in the case of Theorem 4.3 is automatically fulfilled because of the hypothesis* $\zeta_0 < -\alpha - c_1 - 1$*, and in Theorems 4.5 and 4.6 because of the hypothesis* $\epsilon - \alpha - c_1 + 1 \leq n$*. In fact* $n - \alpha$ *is greater than* ϵ*. But this implies that* $\epsilon - n + \alpha < 0$ *and so also the second condition is fulfilled, at least when* $\epsilon \geq 0$*. For the case* $\epsilon < 0$ *in positive characteristic see [16].*

Observe that, if $\epsilon < 0$*, Kodaira, and so* (**C2**)*, holds for every* n*.*

For a general discussion, also in characteristic $p > 0$*, of this question, see section 7, Remark 7.8.*

Remark 4.12. *In the above theorems we assume that $\mathcal{E}$ is a non-split bundle. If $\mathcal{E}$ splits, then (see section 2) $\delta = 0$. In Theorem 4.3 this implies $h^1(\mathcal{E}(n)) - h^2(\mathcal{E}(n)) = 0$ and so nothing can be said on the non-vanishing.*

Let us now consider Theorem 4.6. If $\delta = 0$, then we must have: $\zeta_0 < n < \zeta_0 + \sqrt{-\frac{\tau}{2d} + \frac{\epsilon^2}{4} - \frac{3c_1^2}{4}} \leq \zeta_0 + 1$ (the last inequality depending on Proposition 3.2(8) and (9)). As a consequence ζ_0 cannot be a whole number. Moreover, since we have $2\zeta_0 - \alpha + 1 \leq n < \zeta_0 + \sqrt{-\frac{\tau}{2d} + \frac{\epsilon^2}{4} - \frac{3c_1^2}{4}}$, we obtain that $\zeta_0 < \alpha \leq 0$, hence $\epsilon - c_1 \leq -1$. If $c_1 = 0$, $\epsilon \in \{-1, -3\}$. If $\epsilon = -3$, then n must satisfy the following inequalities: $-\frac{3}{2} < n < -1$ (see Proposition 3.2(8)), which is a contradiction. If $\epsilon = -1$, then, by Proposition 3.2(8), we have $-1 + \alpha + 1 < -\frac{1}{2} + \frac{1}{2} = 0$, which implies $\alpha > 0$, a contradiction. If $c_1 = -1$, then $\epsilon \in \{-2, -4\}$. If $\epsilon = -4$, we have $\sqrt{-\frac{\tau}{2d} + \frac{\epsilon^2}{4} - \frac{3c_1^2}{4}} = \frac{1}{2}$, and so we must have: $-\frac{3}{2} < n < -1$, which is impossible. If $\epsilon = -2$, then $\zeta_0 = -\frac{1}{2}$ and so $-2 - \alpha + 2 < -\frac{1}{2} + \sqrt{1 - \frac{3}{4}}$, which implies $-\alpha < 0$, hence $\alpha > 0$, a contradiction with the non-stability of $\mathcal{E}$.

Then we consider Theorem 4.5. The vanishing of δ on the one hand implies $\lambda > 0$ and $X_0 = 0$. But on the other hand from our hypothesis on the range of n we see that $\zeta_0 \leq -2$, hence $\epsilon = -4, c_1 = 0$. But this contradicts Proposition 3.2(2).

5. Stable Vector Bundles

We start with the following lemma which holds both in the stable and in the non-stable case but is useful only in the present section.

Lemma 5.1. *If $h^1(\mathcal{E}(m)) = 0$ for some integer $m \leq \alpha - 2$, then $h^1(\mathcal{E}(n)) = 0$ for all $n \leq m$.*

Proof. First of all observe that, by our condition (**C3**), from the restriction exact sequence we can obtain in cohomology the exact sequence

$$0 \to H^0(\mathcal{E}(m)) \to H^0(\mathcal{E}(m+1)) \to H^0(\mathcal{E}_H(m+1)) \to 0.$$

Since $m+1 \leq \alpha - 1$ we obtain that $h^0(\mathcal{E}_H(m+1)) = 0$, and so $h^0(\mathcal{E}_H(t)) = 0$ for every $t \leq m+1$. This implies that $h^1(\mathcal{E}(t-1)) \leq h^1(\mathcal{E}(t))$ for each $t \leq m+1$, and so we prove the claim. (Our proof is quite similar to the one given in [17] for $\mathbb{P}^3$, where condition (**C3**) is automatically fulfilled). □

In the present section we assume that $\alpha \geq \frac{\epsilon - c_1 + 5}{2}$, or equivalently that $c_1 + 2\alpha \geq \epsilon + 5$. This means that $\alpha \geq 1$ in any event, so $\mathcal{E}$ is stable.

Theorem 5.2. *Let $\mathcal{E}$ be a rank 2 vector bundle on the threefold X with first relevant level α. If $\alpha \geq \frac{\epsilon + 5 - c_1}{2}$, then $h^1(\mathcal{E}(n)) \neq 0$ for $w_0 \leq n \leq \alpha - 2$.*

Proof. By the hypothesis it holds $w_0 \leq \alpha - 2$, so we have $h^0(\mathcal{E}(n)) = 0$ for all $n \leq w_0 + 1$. Assume $h^1(\mathcal{E}(w_0)) = 0$, then by Lemma 5.1 it holds $h^1(\mathcal{E}(n)) = 0$ for every $n \leq w_0$. Therefore we have

$$\chi(\mathcal{E}(w_0)) = h^0(\mathcal{E}(w_0)) + h^1(\mathcal{E}(-w_0 + \epsilon - c_1)) - h^0(\mathcal{E}(-w_0 + \epsilon - c_1)) = 0.$$

Now observe that the characteristic function has at most three real roots, that are symmetric with respect to ζ_0. Therefore, if w_0 is a root, then $w_0 = \zeta_0 + \sqrt{\vartheta}$ and the other roots are ζ_0 and $\zeta_0 - \sqrt{\vartheta}$. This implies that $\chi(\mathcal{E}(w_0 + 1)) > 0$. On the other hand

$$\chi(\mathcal{E}(w_0 + 1)) = -h^1(\mathcal{E}(w_0 + 1)) \leq 0,$$

a contradiction. So we must have $h^1(\mathcal{E}(w_0)) \neq 0$, then by Lemma 5.1 we obtain the thesis. □

Remark 5.3. *If $\mathcal{E}$ is* ACM, *then $\alpha < \frac{\epsilon+5-c_1}{2}$.*

Theorem 5.4. *Let $\mathcal{E}$ be a normalized rank 2 vector bundle on the threefold X with $\vartheta \geq 0$ and $w_0 < \zeta$. Then the following hold:*

1) *$h^1(\mathcal{E}(n)) \neq 0$ for $\zeta_0 < n < \zeta$, i.e. for $w_0 \leq n \leq \bar{\alpha} - 2$, and also for $n = \bar{\alpha} - 1$ if $\zeta \notin \mathbb{Z}$.*
2) *If $\zeta \in \mathbb{Z}$ and $\alpha < \bar{\alpha}$, then $h^1(\mathcal{E}(\bar{\alpha} - 1)) \neq 0$.*

Proof.

1) The Hilbert polynomial of the bundle $\mathcal{E}$ is strictly negative for each integer such that $w_0 \leq n < \zeta$, but for such an integer n we have $h^2(\mathcal{E}(n)) \geq 0$ and $h^0(\mathcal{E}(n)) - h^0(\mathcal{E}(-n + \epsilon - c_1)) \geq 0$ since $n \geq -n + \epsilon - c_1$ for every $n \geq w_0$, therefore we must have $h^1(\mathcal{E}(n)) \neq 0$. The other statements hold because $\bar{\alpha}$ is, by definition, the integral part of $\zeta + 1$.
2) If $\zeta \in \mathbb{Z}$, then $\zeta = \bar{\alpha} - 1$, so we have $\chi(\mathcal{E}(\bar{\alpha} - 1)) = \chi(\mathcal{E}(\zeta)) = 0$. Moreover $h^0(\mathcal{E}(\bar{\alpha} - 1)) \neq 0$ since $\alpha < \bar{\alpha}$, therefore $h^0(\mathcal{E}(\bar{\alpha} - 1)) - h^3(\mathcal{E}(\bar{\alpha} - 1)) > 0$, and $h^1(\mathcal{E}(n)) = 0$ implies $h^1(\mathcal{E}(m))$, for all $m \leq n$; hence we must have $h^1(\mathcal{E}(\bar{\alpha} - 1)) \neq 0$ to obtain the vanishing of $\chi(\mathcal{E}(\bar{\alpha} - 1))$. □

Remark 5.5. *Observe that in this section we assume $\alpha \geq \frac{\epsilon - c_1 + 5}{2}$, in order to have $w_0 \leq \alpha - 2$ and so to have a non-empty range for n in Theorem 5.2.*

Remark 5.6. *Observe that in the stable case we need not assume any vanishing of $h^1(\mathcal{O}_X(n))$.*

Remark 5.7. *Observe that split bundles are excluded in this section because they cannot be stable.*

6. Examples

We need the following

REMARK 6.1. *Let $X \subset \mathbb{P}^4$ be a smooth threefold of degree d and let f be the projection onto $\mathbb{P}^3$ from a general point of $\mathbb{P}^4$ not on X, and consider a normalized rank two vector bundle $\mathcal{E}$ on $\mathbb{P}^3$ which gives rise to the pull-back $\mathcal{F} = f^*(\mathcal{E})$. We want to check that $f_*(\mathcal{O}_X) \cong \oplus_{i=0}^{d-1}\mathcal{O}_{\mathbb{P}^3}(-i)$.*

Since f is flat and $\deg(f) = d$, $f_*(\mathcal{O}_X)$ is a rank d vector bundle. The projection formula and the cohomology of the hypersurface X shows that $f_*(\mathcal{O}_X)$ is ACM. Thus there are integers $a_0 \geq \cdots \geq a_{d-1}$ such that $f_*(\mathcal{O}_X) \cong \oplus_{i=0}^{d-1}\mathcal{O}_{\mathbb{P}^3}(a_i)$. Since $h^0(X, \mathcal{O}_X) = 1$, the projection formula gives $a_0 = 0$ and $a_i < 0$ for all $i > 0$. Since $h^0(X, \mathcal{O}_X(1)) = 5 = h^0(\mathbb{P}^3, \mathcal{O}_{\mathbb{P}^3}(1)) + h^0(\mathbb{P}^3, \mathcal{O}_{\mathbb{P}^3})$, the projection formula gives $a_1 = -1$ and $a_i \leq -2$ for all $i \geq 2$. Fix an integer $t \leq d-2$ and assume proved $a_i = -i$ for all $i \leq t$ and $a_i < -t$ for all $i > t$. Since $h^0(X, \mathcal{O}_X(t+1)) = \binom{t+5}{4} = \sum_{i=0}^{t}\binom{t+4-i}{3}$, we get $a_{t+1} = -t-1$ and, if $t+1 \leq d-2$, $a_i < -t-1$ for all $i > t+1$. Since $f_*(\mathcal{O}_X) \cong \oplus_{i=0}^{d-1}\mathcal{O}_{\mathbb{P}^3}(-i)$, the projection formula gives the following formula for the first cohomology module:

$$H^i(\mathcal{F}(n)) \cong H^i(\mathcal{E}(n)) \oplus H^i(\mathcal{E}(n-1)) \oplus \cdots \oplus H^i(\mathcal{E}(n-d+1))$$

all i. Observe that, as a consequence of the above equalitiy for $i = 0$, we obtain that $\mathcal{F}$ has the same α as $\mathcal{E}$. Moreover the pull-back $\mathcal{F} = f^*(\mathcal{E})$ and $\mathcal{E}$ have the same Chern class c_1, while $c_2(\mathcal{F}) = d\, c_2(\mathcal{E})$ and therefore $\delta(\mathcal{F}) = d\, \delta(\mathcal{E})$.

Examples:

1. (a stable vector bundle with $c_1 = 0$, $c_2 = 4$ on a quadric hypersurface X).

   Choose $d = 2$ and take the pull-back $\mathcal{F}$ of the stable vector bundle $\mathcal{E}$ on $\mathbb{P}^3$ of [17], example 4.1. Then the numbers of $\mathcal{F}$ (see Notation) are: $c_1 = 0$, $c_2 = 4$, $\alpha = 1$, $\bar{\alpha} = 2$, $\zeta_0 = -\frac{3}{2}$, $w_0 = -1$, $\vartheta = \frac{25}{4}$, $\zeta = -\frac{3}{2} + \sqrt{\frac{25}{4}} = 1 \in \mathbb{Z}$. From [17], example 4.1, we know that $h^1(\mathcal{E}) \neq 0$. Since $H^1(\mathcal{F}(1)) \cong H^1(\mathcal{E}(1)) \oplus H^1(\mathcal{E})$, we have: $h^1(\mathcal{F}(1)) \neq 0$, one shift higher than it is stated in Theorem 5.4(2).

2. (a non-stable vector bundle with $c_1 = 0$, $c_2 = 45$ on a hypersurface of degree 5).

   Choose $d = 5$ and take the pull-back $\mathcal{F}$ of the stable vector bundle $\mathcal{E}$ on $\mathbb{P}^3$ of [17], example 4.5. Then the numbers of $\mathcal{F}$ (see Notation) are: $c_1 = 0$, $c_2 = 45$, $\alpha = -3$, $\delta = 90$, $\zeta_0 = 0$. From [17], Theorem 3.8, we know that $h^1(\mathcal{E}(12)) \neq 0$. Since $H^1(\mathcal{F}(16)) \cong H^1(\mathcal{E}(16)) \oplus \cdots \oplus H^1(\mathcal{E}(12))$, we have: $h^1(\mathcal{F}(16)) \neq 0$ (Theorem 4.5 states that $h^1(\mathcal{F}(10)) \neq 0$.

3. (a stable vector bundle with $c_1 = -1$, $c_2 = 2$ on a quadric hypersurface).

   Let $\mathcal{E}$ be the rank two vector bundle corresponding to the union of two skew lines on a smooth quadric hypersurface $Q \subset \mathbb{P}^4$. Then its numbers are : $c_1 = -1$, $c_2 = 2$, $\alpha = 1$ and it is known that $h^1(\mathcal{E}(n)) \neq 0$ if and only if $n = 0$.

   Observe that in this case $\vartheta = \frac{5}{2} \geq 0$, $\zeta_0 = -1$, $\bar{\alpha} = 1$. Therefore Theorem 5.4 states exactly that $h^1(\mathcal{E}) \neq 0$, hence this example is sharp.

4. (a non-stable vector bundle with $c_1 = 0$, $c_2 = 8$ on a quadric hypersurface).

   Choose $d = 2$ and take the pull-back $\mathcal{F}$ of the non-stable vector bundle $\mathcal{E}$ on $\mathbb{P}^3$ of [17], example 4.10. Then the numbers of $\mathcal{F}$ (see Notation) are: $c_1 = 0$, $c_2 = 8$, $\alpha = 0$, $\zeta_0 = -\frac{3}{2}$, $\delta = 8$. We know (see [17], example 4.10) that $h^1(\mathcal{E}(2)) \neq 0, h^1(\mathcal{E}(3)) = 0$. Since $H^1(\mathcal{F}(3)) \cong H^1(\mathcal{E}(3)) \oplus H^1(\mathcal{E}(2))$, we have: $h^1(\mathcal{F}(3)) \neq 0$, exactly the bound of Theorem 4.6.

Remark 6.2. *The bounds for a degree d threefold in $\mathbb{P}^4$ agree with [17], where $\mathbb{P}^3$ is considered.*

7. Threefolds with $\mathrm{Pic}(X) \neq \mathbb{Z}$

Let X be a smooth and connected projective threefold defined over an algebraically closed field $\mathbf{k}$. Let $\mathrm{Num}(X)$ denote the quotient of $\mathrm{Pic}(X)$ by numerical equivalence. Numerical classes are denoted by square brackets [ ]. We assume $\mathrm{Num}(X) \cong \mathbb{Z}$ and take the unique isomorphism $\eta\colon \mathrm{Num}(X) \to \mathbb{Z}$ such that 1 is the image of a fixed ample line bundle. Notice that $M \in \mathrm{Pic}(X)$ is ample if and only if $\eta([M]) > 0$.

Remark 7.1. *Let $\eta\colon \mathrm{Num}(X) \to \mathbb{Z}$ be as before. Notice that every effective divisor on X is ample and hence its η is strictly positive. For any $t \in \mathbb{Z}$ set $\mathrm{Pic}_t(X) := \{L \in \mathrm{Pic}(X) \mid \eta([L]) = t\}$. Hence $\mathrm{Pic}_0(X)$ is the set of all isomorphism classes of numerically trivial line bundles on X. The set $\mathrm{Pic}_0(X)$ is parametrized by a scheme of finite type ([10], Proposition 1.4.37). Hence for each $t \in \mathbb{Z}$ the set $\mathrm{Pic}_t(X)$ is bounded. Let now $\mathcal{E}$ be a rank 2 vector bundle on X. Since $\mathrm{Pic}_1(X)$ is bounded there is a minimal integer t such that there is $B \in \mathrm{Pic}_t(X)$ and $h^0(\mathcal{E} \otimes B) > 0$. Call it $\alpha(\mathcal{E})$ or just α. By the definition of α there is $B \in \mathrm{Pic}_\alpha(X)$ such that $h^0(X, \mathcal{E} \otimes B) > 0$. Hence there is a non-zero map $j\colon B^* \to \mathcal{E}$. Since B^* is a line bundle and $j \neq 0$, j is injective. The definition of α gives the non-existence of a non-zero effective divisor D such that j factors through an inclusion $B^* \to B^*(D)$, because $\eta([D]) > 0$. Thus the inclusion j induces an exact sequence*

$$0 \to B^* \to \mathcal{E} \to \mathcal{I}_Z \otimes B \otimes \det\ (\mathcal{E}) \to 0 \tag{3}$$

in which Z is a closed subscheme of X of pure codimension 2.

Observe that $\eta([B]) = \alpha, \eta([B^*]) = -\alpha, \eta([B \otimes det(\mathcal{E})]) = \alpha + c_1$, hence the exact sequence is quite similar to the usual exact sequence that holds true in the case $\mathrm{Pic}(X) \cong \mathbb{Z}$.

Notation. We set $\epsilon := \eta([\omega_X])$, $\alpha := \alpha\ (\mathcal{E})$ and $c_1 := \eta([\det\ (\mathcal{E})])$. So we can speak of a normalized vector bundle $\mathcal{E}$, with $c_1 \in \{0, -1\}$. Moreover we say that $\mathcal{E}$ is stable if $\alpha > 0$, non-stable if $\alpha \le 0$. Furthermore ζ_0, ζ, w_0, $\bar{\alpha}$, ϑ are defined as in section 2.

Remark 7.2. *Fix any $L \in \mathrm{Pic}_1(X)$ and set: $d = L^3 =$ degree of X. The degree d does not depend on the numerical equivalence class. In fact, if R is numerically equivalent to 0, then $(L+R)^3 = L^3+R^3+3L^2R+3LR^2 = L^3+0+0+0 = L^3$. Then it is easy to see that the formulas for $\chi(\mathcal{O}_X(n))$ and $\chi(\mathcal{E}(n))$ given in section 2 still hold if we consider $\mathcal{O}_X \otimes L^{\otimes n}$ and $\mathcal{E} \otimes L^{\otimes n}$ (see [18]).*

Remark 7.3.

(a) *Assume the existence of $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = 1$ and $h^0(X, L) > 0$. Then for every integer $t > \alpha$ there is $M \in \mathrm{Pic}(X)$ such that $\eta([M]) = t$ and $h^0(X, \mathcal{E} \otimes M) > 0$.*

(b) *Assume $h^0(X, L) > 0$ for every $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = 1$. Then $h^0(X, \mathcal{E} \otimes M) > 0$ for every $M \in \mathrm{Pic}(X)$ such that $\eta([M]) > \alpha$.*

Proposition 7.4. *Let $\mathcal{E}$ be a normalized rank two vector bundle and assume the existence of a spanned $R \in \mathrm{Pic}(X)$ such that $\eta([R]) = 1$. If $\mathrm{char}(\mathbf{k}) > 0$, assume that $|R|$ induces an embedding of X outside finitely many points. Assume*

$$2\alpha \le -\epsilon - 3 - c_1 \tag{4}$$

and $h^1(X, \mathcal{E} \otimes N) = 0$ for every $N \in \mathrm{Pic}(X)$ such that $\eta([N]) \in \{-\alpha - c_1 - 1, \alpha+2+e\}$. If $h^1(X, B) = 0$ for every $B \in \mathrm{Pic}(X)$ such that $\eta([B]) = -2\alpha - c_1$, then $\mathcal{E}$ splits.

If moreover $h^1(X, M) = 0$ for every $M \in \mathrm{Pic}(X)$ then it is enough to assume that $h^1(X, \mathcal{E} \otimes N) = 0$ for every $N \in \mathrm{Pic}(X)$ such that $\eta([N]) = -\alpha - c_1 - 1$.

Proof. By assumption there is $M \in \mathrm{Pic}(X)$ such that $\eta([M]) = \alpha$ and $h^0(X, \mathcal{E} \otimes M) > 0$. Set $A := M^*$. We have seen in remark 7.1 that $\mathcal{E}$ fits into an extension of the following type:

$$0 \to A \to \mathcal{E} \to \mathcal{I}_C \otimes \det(\mathcal{E}) \otimes A^* \to 0 \tag{5}$$

with C a locally complete intersection closed subscheme of pure dimension 1. Let H be a general element of $|R|$ and T the intersection of H with another general element of $|R|$. Observe that T, under our assumptions, is generically reduced by Bertini's Theorem (see [6], Theorem II, 8.18 and Remark II, 8.18.1). Since R is spanned, T is a locally complete intersection curve and $C \cap T = \emptyset$. Hence $\mathcal{E}|_T$ is an extension of $\det(\mathcal{E}) \otimes A^*|_T$ by $A|_T$. Since T is generically reduced and locally a complete intersection, it is reduced. Hence $h^0(T, M^*) = 0$ for every ample line bundle M on T. Since $\omega_T \cong (\omega_X \otimes R^{\otimes 2})|_T$, we have $\dim(\mathrm{Ext}^1_T(\det(\mathcal{E}) \otimes A^*, A)) = h^0(T, (\det(\mathcal{E}) \otimes (A^*)^{\otimes 2} \otimes \omega_X \otimes R^{\otimes 2})|_T) = 0$ (indeed $\eta([\det(\mathcal{E}) \otimes (A^*)^{\otimes 2} \otimes \omega_X \otimes R^{\otimes 2}]) = 2\alpha + c_1 + e + 2 < 0$). Hence $\mathcal{E}|_T \cong A|_T \oplus (\det(\mathcal{E}) \otimes A^*)|_T$. Let σ be the non-zero section of $(\mathcal{E} \otimes (A \otimes \det(\mathcal{E})^*)|_T$ coming from the projection onto the second factor of the decomposition just given. The vector bundle $\mathcal{E}|_H$ is an extension of $(\det(\mathcal{E}) \otimes A^*)|_H$ by $A|_H$ if and only if $C \cap H = \emptyset$. Since R is ample, $C \cap H = \emptyset$ if and only if $C = \emptyset$. Hence we get simultaneously $C \cap H = \emptyset$ and $\mathcal{E}|_H \cong A|_H \oplus (\det(\mathcal{E}) \otimes A^*)|_H$ if we prove the existence of $\tau \in H^0(H, (\mathcal{E} \otimes (A \otimes \det(\mathcal{E})^*)|_H)$ such that $\tau|_T = \sigma$. To get τ it is sufficient to have $H^1(H, (E \otimes (A \otimes \det(\mathcal{E})^* \otimes R^*)|_H) = 0$. A standard exact sequence shows that $H^1(H, (\mathcal{E} \otimes (A \otimes \det(\mathcal{E})^* \otimes R^*)|_H) = 0$ if $h^1(X, (\mathcal{E} \otimes (A \otimes \det(\mathcal{E})^* \otimes R^*)) = 0$ and $h^2(X, (\mathcal{E} \otimes (A \otimes \det(\mathcal{E})^* \otimes R^* \otimes R^*)) = 0$. Since $\mathcal{E}^* \cong \mathcal{E} \otimes \det(\mathcal{E})^*$, Serre duality gives $h^2(X, \mathcal{E} \otimes (A \otimes \det(\mathcal{E})^* \otimes R^* \otimes R^*)) = h^1(X, \mathcal{E} \otimes A \otimes R^{\otimes 2} \otimes \omega_X)$. Since $\eta([A \otimes \det(\mathcal{E})^* \otimes R^*]) = -\alpha - c_1 - 1$ and $\eta([A \otimes R^{\otimes 2} \otimes \omega_X]) = \alpha + e + 2$, we get that $C = \emptyset$. The last sentence follows because $\eta([A^{\otimes 2} \otimes \det(\mathcal{E})^*]) = -2\alpha - c_1$. □

REMARK 7.5. *Fix integers $t < z \leq \alpha - 2$. Assume the existence of $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = z$ and $h^1(X, \mathcal{E} \otimes L) = 0$. If there is $R \in \mathrm{Pic}(X)$ such that $\eta([R]) = 1$ and $h^0(X, R) > 0$, then there exists $M \in \mathrm{Pic}(X)$ such that $\eta([M]) = t$ and $h^1(X, \mathcal{E} \otimes M) = 0$. If $h^0(X, R) > 0$ for every $R \in \mathrm{Pic}(X)$ such that $\eta([R]) = 1$, then $h^1(X, \mathcal{E} \otimes M) = 0$ for every $M \in \mathrm{Pic}(X)$ such that $\eta([M]) = t$.*

The proof can follow the lines of Lemma 5.1. In fact consider a line bundle R with $\eta([R]) = 1$ and let H be the zero-locus of a non-zero section of R; then we have the following exact sequence:

$$0 \to \mathcal{E} \otimes L \to \mathcal{E} \otimes L \otimes R \to (\mathcal{E} \otimes L \otimes R)|_H \to 0.$$

Now observe that the vanishing of $h^1(X, \mathcal{E} \otimes L)$ implies that $h^0((\mathcal{E} \otimes L \otimes R)|_H) = 0$. And now we can argue as in Lemma 5.1 (see also [17]).

REMARK 7.6.

(a) *Assume the existence of $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = 1$ and $h^0(X, L) > 0$. Then for every integer $t > \alpha$ there is $M \in \mathrm{Pic}(X)$ such that $\eta([M]) = t$ and $h^0(X, \mathcal{E} \otimes M) > 0$.*

(b) *Assume $h^0(X, L) > 0$ for every $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = 1$. Then $h^0(X, \mathcal{E} \otimes M) > 0$ for every $M \in \mathrm{Pic}(X)$ such that $\eta([M]) > \alpha$.*

Remark 7.7. *In all our results of sections 4 and 5 we use the vanishing of* $h^1(\mathcal{O}_X(n))$ *for all* n *(and by Serre duality of* $h^2(\mathcal{O}_X(n))$*) (or, at least,* $\forall n \notin \{0, \cdots, \epsilon\}$*), see Remark 4.11.*

From now on we need to use similar vanishing conditions and so we introduce the following condition:

> (**C4**) $h^1(X, L) = 0$ *for all* $L \in \mathrm{Pic}(X)$ *such that either* $\eta([L]) < 0$ *or* $\eta([L]) > \epsilon$.

Observe that (**C4**) *is always satisfied in characteristic 0 (by the Kodaira vanishing theorem). In positive characteristic it is often satisfied. This is always the case if* X *is an abelian variety ([12] page 150).*

Observe also that, if $\epsilon \leq -1$*, the Kodaira vanishing and our condition put no restriction on* n *(see also Remark 4.12).*

Example. If (4) holds, then $-2\alpha - c_1 > \epsilon$. Hence we may apply Proposition 7.4 to X. In particular observe that, in the case of an abelian variety with $\mathrm{Num}(X) \cong \mathbb{Z}$ or in the case of a Calabi-Yau threefold with $\mathrm{Num}(X) \cong \mathbb{Z}$, we have $\epsilon = 0$. Notice that Proposition 7.4 also applies to any threefold X whose ω_X has finite order.

With the assumption of condition (**C4**) the proofs of Theorems 4.3, 4.5, 4.6 can be easily modified in order to obtain the statements below ($\mathcal{E}$ is normalized, i.e. $\eta([\det(\mathcal{E})]) \in \{-1, 0\}$), where, by the sake of simplicity, we assume $\epsilon \geq 0$ (if $\epsilon < 0$, (**C4**), which holds by [16], implies that all the vanishing of h^1 and h^2 for all $L \in \mathrm{Pic}(X)$ hold).

Theorem 7.8. *Assume* (**C4**), $\alpha \leq 0$, *the existence of* $R \in \mathrm{Pic}(X)$ *such that* $\eta([R]) = 1$ *and* $\zeta_0 < -\alpha - c_1 - 1$. *Fix an integer* n *such that* $\zeta_0 < n \leq -\alpha - 1 - c_1$. *Fix* $L \in \mathrm{Pic}(X)$ *such that* $\eta([L]) = n$. *Then* $h^1(\mathcal{E} \otimes L) \geq (n - \zeta_0)\delta > 0$.

Remark 7.9. *Observe that we should require the following conditions:* $n - \alpha \notin \{0, \ldots, \epsilon\}, \epsilon - n + \alpha \notin \{0, \ldots, \epsilon\}$. *But they are automatically fulfiled under the assumption that* $\zeta_0 < -\alpha - c_1 - 1$.

Theorem 7.10. *Assume* (**C4**), $\alpha \leq 0$, *the existence of* $R \in \mathrm{Pic}(X)$ *such that* $\eta([R]) = 1$ *and the same hypotheses of Theorem 4.6. Fix* $L \in \mathrm{Pic}(X)$ *such that* $\eta([L]) = n$. *Then* $h^1(\mathcal{E} \otimes L) \geq -S(n) > 0$ *(*$S(n)$ *being defined as in Theorem 4.6).*

Theorem 7.11. *Assumption as in Theorem 4.5. Moreover assume* (**C4**) *and* $n - \alpha \notin \{0, \ldots, \epsilon\}$. *Fix* $L \in \mathrm{Pic}(X)$ *such that* $\eta([L]) = n$. *Then* $h^1(\mathcal{E} \otimes L) \geq -\frac{d}{6}F(n + \alpha - \zeta_0 + \frac{c_1}{2}) > 0$ *(*F *being defined as in Theorem 4.5).*

Remark 7.12. *Observe that in Theorems 7.10 and 7.11 we should require* $n - \alpha \notin \{0, \ldots, \epsilon\}$, *but the assumption* $\epsilon - \alpha - c_1 + 1 \leq n$ *implies that it is automatically fulfilled.*

The proofs of the above theorems are based on the existence of the exact sequence (3) and on the properties of α. They follow the lines of the proofs given in the case $\mathrm{Pic}(X) \cong \mathbb{Z}$. Here and in section 4 we actually need only the Kodaira vanishing (true in characteristic 0 and assumed in characteristic $p > 0$) and no further vanishing of the first cohomology.

Also the stable case can be extended to a smooth threefold with $\mathrm{Num}(X) \cong \mathbb{Z}$. Observe that the proofs can follow the lines of the proofs given in the case $\mathrm{Pic}(X) \cong \mathbb{Z}$ and make use of Remark 7.6 (which extends Theorem 5.1).

More precisely we have:

THEOREM 7.13. *Assumptions as in Theorem 5.2 and fix $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = n$. Then, if $\alpha \geq \frac{\epsilon+5-c_1}{2}$, then $h^1(\mathcal{E} \otimes L) \neq 0$ for $w_0 \leq n \leq \alpha - 2$.*

THEOREM 7.14. *Assumptions as in Theorem 5.4 and fix $L \in \mathrm{Pic}(X)$ such that $\eta([L]) = n$. Then the following hold:*

1) *$h^1(\mathcal{E} \otimes L) \neq 0$ for $\zeta_0 < n < \zeta$, i.e. for $w_0 \leq n \leq \bar{\alpha} - 2$, and also for $n = \bar{\alpha} - 1$ if $\zeta \notin \mathbb{Z}$.*

2) *If $\zeta \in \mathbb{Z}$ and $\alpha < \bar{\alpha}$, then $h^1(\mathcal{E} \otimes N) \neq 0$, for every N such that $\eta([N]) = \bar{\alpha} - 1$.*

REMARK 7.15. *The above theorems can be applied to any X such that $\mathrm{Num}(X) \cong \mathbb{Z}$, $\epsilon = 0$ and $h^1(X, L) = 0$ for all $L \in \mathrm{Pic}(X)$ such that $\eta([L]) \neq 0$, for instance to $X =$ an abelian threefold with $\mathrm{Num}(X) \cong \mathbb{Z}$.*

REMARK 7.16. *If X is any threefold (in characteristic 0 or positive) such that $h^1(X, L) = 0$, for all $L \in \mathrm{Pic}(X)$, then we can avoid the restriction $n - \alpha \notin \{0, \ldots, \epsilon\}$. Not many threefolds, beside any $X \subset \mathbb{P}^4$, fulfill these conditions.*

REMARK 7.17. *Observe that in Theorems 7.13 and 7.14 we do not assume* (**C4**) *(see also Remark 5.6).*

REMARK 7.18. *Observe that also in the present case* ($\mathrm{Num}(X) \cong \mathbb{Z}$)*, we have: $\delta = 0$ if and only if $\mathcal{E}$ splits. Therefore Remarks 4.12 and 5.7 apply here.*

REFERENCES

[1] L. CHIANTINI AND C. MADONNA, *A splitting criterion for rank* 2 *bundles on a general sextic threefold*, Int. J. Math. **15** (2004), 341–359.

[2] L. CHIANTINI AND P. VALABREGA, *Subcanonical curves and complete intersections in projective* 3*-space*, Ann. Mat. Pura Appl. **138** (1984), 309–330.

[3] L. CHIANTINI AND P. VALABREGA, *On some properties of subcanonical curves and unstable bundles*, Comm. Algebra **15** (1987), 1877–1887.

[4] PH. ELLIA, *Sur la cohomologie de certains fibrés de rang deux sur* $\mathbb{P}^3$, Ann. Univ. Ferrara **38** (1992), 217–227.

[5] G. Gherardelli, *Sulle curve sghembe algebriche intersezioni complete di due superficie*, Atti Reale Accad. Italia **4** (1943), 128–132.

[6] R. Hartshorne, *Algebraic geometry*, Graduate texts in mathematics volume 52, Springer, Berlin (1977).

[7] R. Hartshorne, *Stable vector bundles of rank* 2 *on* $\mathbb{P}^3$, Math. Ann. **238** (1978), 229–280.

[8] R. Hartshorne, *Stable reflexive sheaves*, Math. Ann. **254** (1980), 121–176.

[9] R. Hartshorne, *Stable reflexive sheaves II*, Invent. Math. **66** (1982), 165–190.

[10] R. Lazarsfeld, *Positivity in algebraic geometry I*, Springer, Berlin (2004).

[11] C. Madonna, *A splitting criterion for rank* 2 *vector bundles on hypersurfaces in* $\mathbb{P}^4$, Rend. Sem. Mat. Univ. Pol. Torino **56** (1998), 43–54.

[12] D. Mumford, *Abelian varieties*, Oxford University Press, Oxford (1974).

[13] S. Popescu, *On the splitting criterion of Chiantini and Valabrega*, Rev. Roumaine Math. Pures Appl. **33** (1988), 883–887.

[14] G.V. Ravindra, *Curves on threefolds and a conjecture of Griffiths-Harris*, Math. Ann. **345** (2009), 731–748.

[15] M. Roggero and P. Valabrega, *Some vanishing properties of the intermediate cohomology of a reflexive sheaf on* $\mathbb{P}^n$, J. Algebra **170** (1994), 307–321.

[16] N. Shepherd-Barron, *Fano threefolds in positive characteristic*, Comp. Math. **105** (1997), 237–265.

[17] P. Valabrega and M. Valenzano, *Non-vanishing theorems for non-split rank* 2 *bundles on* $\mathbb{P}^3$*: a simple approach*, Atti Acc. Peloritana **87** (2009), 1–18.

[18] M. Valenzano, *Rank* 2 *reflexive sheaves on a smooth threefold*, Rend. Sem. Mat. Univ. Pol. Torino **62** (2004), 235–254.

Authors' addresses:

Edoardo Ballico
Dipartimento di Matematica
Università di Trento
Via Sommarive 14, 38123 Povo TN, Italy
E-mail: `ballico@science.unitn.it`

Paolo Valabrega
Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
E-mail: `paolo.valabrega@polito.it`

Mario Valenzano
Dipartimento di Matematica
Università di Torino
Via Carlo Alberto 10, 10123 Torino, Italy
E-mail: `mario.valenzano@unito.it`

Received October 25, 2010
Revised March 5, 2011

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 31–51

A Proof of Monge Problem in $\mathbb{R}^n$ by Stability

Laura Caravenna

Abstract. *The Monge problem in $\mathbb{R}^n$, with a possibly asymmetric norm cost function and absolutely continuous first marginal, is generally underdetermined. An optimal transport plan is selected by a secondary variational problem, from a work on crystalline norms. In this way the mass still moves along lines. The paper provides a quantitative absolute continuity push forward estimate for the translation along these lines: the consequent area formula, for the disintegration of the Lebesgue measure w.r.t. the partition into these $1D$-rays, shows that the conditional measures are absolutely continuous, and yields uniqueness of the optimal secondary transport plan non-decreasing along rays, recovering that it is induced by a map.*

Keywords: Monge Problem, Area Estimates, Disintegration of Measures.
MS Classification 2010: 49J45, 49K30, 28A50, 49Q20

1. Introduction

Topic of this note is a sharp area push forward estimate relative to a solution to the Kantorovich problem in $\mathbb{R}^n$, when the cost function is given by a possibly asymmetric norm $\|\cdot\|$ and only the first marginal is assumed to be absolutely continuous, without assuming strict convexity of the norm. In particular, this provides a proof of existence of solutions to the Monge problem which is based on a 1-dimensional disintegration technique relying on the stability of a particular solution of the problem. Given two Borel probability measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^n)$, we study the minimization of the functional

$$\mathcal{I}_M(t) = \int_{\mathbb{R}^n} \|t(x) - x\| \, d\mu(x) \tag{MP}$$

among the Borel maps $t : \mathbb{R}^n \to \mathbb{R}^n$ whose image measure of μ is ν. We prove that the additional optimality conditions choosen in [2] determine a unique optimal transport map, selected also in [20], under the natural assumption that μ is absolutely continuous w.r.t. the Lebesgue measure $\mathcal{L}^n$. This assumption is necessary, as shown in Section 8 of [3].

The strategy is by reduction to 1-dimensional transport problems. It consists mainly of an area formula for the Lebesgue measure which allows the reduction: we prove a regularity of the disintegration along rays of the limit plan by the one of the approximations — once convergence is established. The limiting procedure is not based on Hopf-Lax formula of potential functions but on a uniqueness criterion. It is a particular case of a more general result in a forthcoming work by Bianchini and Daneri. The technique has been used in [9] in 2007, and then [17], improved simplifying the basic estimate in [7, 18].

Before introducing this work, we present a brief review of the main literature.

1.1. An Account on the Literature

The original Monge problem arose in 1781 for continuous masses μ, ν supported on compact, disjoint sets in dimension 2, 3 and with the cost defined by the Euclidean norm ([31]). Monge himself conjectured important features of the transport, such as, with the Euclidean norm, the facts that two transport rays may intersect only at endpoints and that the directions of the transported particles form a family of normals to some family of surfaces.

Investigated first in [4, 21], the problem was left apart for a long period.

A fundamental improvement in the understanding came with the relaxation of the problem in the space of probability measures ([26, 27]), consisting in the Kantorovich formulation. Instead of looking at maps in $\mathbb{R}^n$, ones considers the following minimization problem in the space $\Pi(\mu, \nu)$ of couplings between μ and ν: minimizing the linear functional

$$\mathcal{I}_K(\pi) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|y - x\| \, d\pi(x, y) \tag{KP}$$

among the *transport plans* π, defined as members of the set

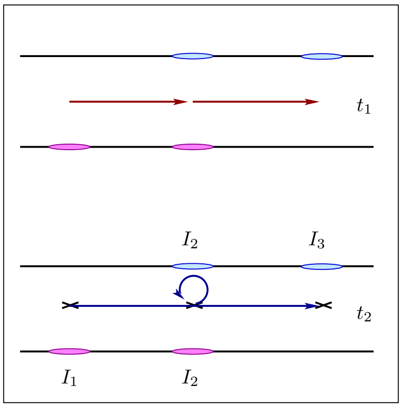
$$\Pi(\mu, \nu) = \{\pi \in \mathcal{M}^+ : \; p^x_\sharp \pi = \mu, \; p^y_\sharp \pi = \nu\},$$

where p^x, p^y are respectively the projections on the first and on the second factor space of $\mathbb{R}^n \times \mathbb{R}^n$. Notice that $\Pi(\mu, \nu)$ is convex, w^*-compact.

In particular, minimizers to (KP) always exist by the direct method of Calculus of Variations. The formulation (KP) is indeed a generalization of the model, allowing that mass at some point can be split to more destinations. Therefore, a priori the minimum value in (MP) is higher than the one in (KP), and the minimizers of the latter are not suitable for the former.

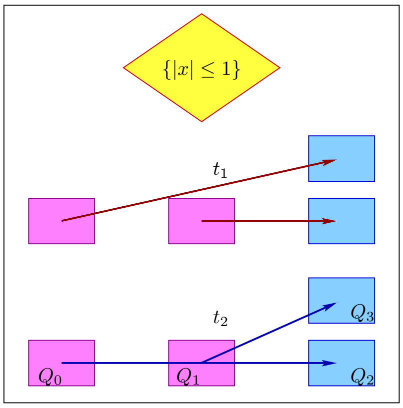
A standard approach to (MP) consists in showing that at least one of the optimizers to (KP) is concentrated on the graph of a function.

This is plainly effective when the cost is given by the squared Euclidean distance instead of $\|y - x\|$: by the uniform convexity there exists a unique



(a) One dimensional example.

Let μ be the Lebesgue measure on $I_1 \cup I_2 \subset \mathbb{R}$ and ν the Lebesgue measure on $I_2 \cup I_3$. Both the maps t_1 translating I_1 to I_2, I_2 to I_3 and the map t_2 translating I_1 to I_3 and leaving I_2 fixed are optimal. Moreover, any convex combination of the two transport plans induced by t_1, t_2 is again a minimizer for (KP), but clearly it is not induced by a map.



(b) Two dimensional example.

The unit ball of $\|\cdot\|$ is given by the rhombus. Let μ be the Lebesgue measure $Q_0 \cup Q_1 \subset \mathbb{R}^2$ and ν the Lebesgue measure on $Q_2 \cup Q_3$. Both the maps t_1, t_2 translating one of the first two squares to one of the second to squares are optimal, and they transport mass in different directions.

Figure 1: The optimal transport map with a generic norm is not unique.

optimizer π to (KP) of the form $\pi = (\mathrm{Id}, \mathrm{Id} - \nabla\phi)_\sharp \mu$ for a semiconcave function ϕ, the Kantorovich potential. Therefore, when $\mu \ll \mathcal{L}^n$ and $\nu \ll \mathcal{L}^n$, the optimal map is μ-a.e. defined by $x \mapsto x - \nabla\phi(x)$ and it is one-to-one ([12, 13, 28] are the first results, extended to *uniformly convex* functions of the distance e.g. in [32, 30, 25, 15]).

However, even in the case of the Euclidean norm, it is well known that this approach presents difficulties: at $\mathcal{L}^n$-a.e. point the Kantorovich potential fixes the direction of the transport, but not the precise point where the mass goes to. This is a feature of the problem, also in dimension one (see the example in Figure 1a).

The data are not sufficient to determine a single transport map, since there is no uniqueness. Uniqueness can be recovered with the further requirement of monotonicity along transport rays ([24]).

The situation becomes even more complicated with a generic norm cost function, instead of the Euclidean one. The symmetry of the norm plays no role, but the loss in strict convexity of the unit ball is relevant, since the transport may not occur along lines and the direction of the transport can vary (see the example in Figure 1b).

The Euclidean case, and thus the one proposed by Monge, has been rigorously solved only around 2000 in [22, 35, 3, 14].

Roughly, the approaches in the last three papers is at least partially based on a decomposition of the domain into 1-dimensional invariant regions for the transport, called transport rays. Due to the strict convexity of the unit ball, these regions are 1-dimensional convex sets. Due to regularity assumptions on the unit ball and a clever countable partition of the ambient space, it is moreover possible to reduce to the case where the directions of these segments is Lipschitz continuous. This, by Area or Coarea formula, allows to disintegrate the Lebesgue measure w.r.t. the partition in transport rays, obtaining absolutely continuous conditional probabilities on the 1-dimensional rays. In turn, this suffices to perform a reduction argument, that we also use in the present paper, which yields the thesis: indeed, one can fix within each ray an optimal transport map, uniquely defined imposing monotonicity within each ray. However, as in [9, 17, 16], we do not rely on any Lipschitz regularity of the vector field of directions for deriving an Area formula.

This kind of approach was introduced already in 1976 by Sudakov ([34]), in the more generality of a possibly asymmetric norm — which actually is the case we are considering. However, its argument remains incomplete: a regularity property of the disintegration of the Lebesgue measure w.r.t. decompositions of the space into affine regions was not proved correctly, and, actually, stated in a form which does not hold ([1]). Indeed, there exists a compact subset of the unit square having measure 1 and made of disjoint segments, with Borel direction, such that the disintegration of the Lebesgue measure w.r.t. the partition in segments has atomic conditional measures ([29], in [2] improved by Alberti et al.). The reduction argument described above requires instead absolutely continuous conditional measures, in order to solve the 1-dimensional transport problems, and therefore a regularity of the partition in transport rays must be proved. In the case of a strictly convex norm the affine regions reduce to lines and Sudakov argument was completed in [17]. In this paper we follow the alternative 1-dimensional decomposition selected by the additional variational principles, instead of the affine one considered by Sudakov. We choose the selection of [2], chosen also in [20].

The method in [22] is based on PDEs and they introduce the concept of transport density, widely studied since there — the very first works are [23, 1, 11, 24]. In [33] one finds more references as well as summability estimates obtained by interpolation and a limiting procedure of the kind also of this note; these are proved for the Euclidean distance, but they should work as well in this setting. Given a Kantorovich potential u for the transport problem between two absolutely continuous measures with compactly supported and smooth densities f^+, f^-, they define as transport density a nonnegative function a supported on the family of transport rays and satisfying

$$-\operatorname{div}(a\nabla u) = f^+ - f^-$$

in distributional sense. The above equation was present already in [5] with different motivation. It allows a generalization to measures, and an alternative definition introduced first in [10] for $\rho := a\mathcal{L}^n$ is given by the Radon measure defined on $A \in \mathcal{B}(\mathbb{R}^n)$ as

$$\rho(A) := \int_{\mathbb{R}^n \times \mathbb{R}^n} \mathcal{H}^1 \llcorner (A \cap [\![x, y]\!])\, d\pi(x, y), \tag{1}$$

where π is an optimal transport plan.

When the unit ball in not *strictly* convex, the first results available were given in [2] for the 2-dimensional case, completely solved, and for crystalline norms. Their strategy is to fix both the direction of the transport and the transport map by imposing additional optimality conditions, and then to carry out a Sudakov-type argument on the selected transports.

We follow the same strategy, and the disintegration technique from [6, 9].

A different proof of existence for general norms, with a selection based on the same optimality conditions, has been presented in [20], improving their argument for strictly convex norms in [19]. It does not arrive to disintegration of measures, it is more concerned with the regularity of the transport density. Also their argument is based on the geometric constraint that c_{s}-monotonicity impose on c_{s}-optimal transference plans, and an intermediate step is to prove that the set of initial points of secondary rays of a limit plan π, of the same maps we consider, is Lebesgue negligible. This important observation was also used for the solution in the special 2-dimensional case in [2], and generalized in more dimensions in [6, 9].

1.2. Topic of this Paper

By a *possibly asymmetric norm* $\|\cdot\|$ we mean a continuous function $\mathbb{R}^n \to [0, +\infty)$ having convex sublevel sets, containing the origin in the interior, and which is positively homogeneous ($\lambda\|x\| = \|\lambda x\|$ for $\lambda \geq 0$ and $x \in \mathbb{R}^n$). The study of this paper lies in the context of the following general problem, difficult due to the degeneracy and non-smoothness of the norm.

Primary Transport Problem. *Consider the Monge-Kantorovich optimal transport problem*

$$\min_{\pi \in \Pi(\mu,\nu)} \int \|y - x\|\, \mathrm{d}\pi(x, y) \tag{2}$$

between two positive Radon measures μ, ν with the same total variation, assuming that $\mu \ll \mathcal{L}^n$.

In order to avoid triviality we suppose that there exists a transport plan with finite cost. Since there is no uniqueness by the lack of *strict* convexity, one considers the family $\mathcal{O}_{\mathrm{p}} \subset \Pi(\mu,\nu)$ of minimizers to the primary problem. We call the members of $\mathcal{O}_{\mathrm{p}}$ the *optimal primary transport plans*. Let ϕ be a Kantorovich potential for this primary problem, by which we mean a function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\phi(x) - \phi(y) \leq \|y - x\| \qquad \forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \tag{3a}$$

$$\phi(x) - \phi(y) = \|y - x\| \qquad \text{for } \pi\text{-a.e. } (x,y), \quad \forall \pi \in \mathcal{O}_{\mathrm{p}}. \tag{3b}$$

We select then particular minimizers by the following secondary problem.

Secondary Transport Problem. *Consider a strictly convex norm* $|\cdot|$. *Study*

$$\min_{\pi \in \mathcal{O}_{\mathrm{p}}} \int |y - x| \, \mathrm{d}\pi(x,y) = \min_{\pi \in \Pi(\mu,\nu)} \int c_{\mathrm{s}}(x,y) \, d\pi(x,y) \tag{4}$$

where the secondary cost function c_{s} *is defined by*

$$c_{\mathrm{s}}(x,y) := \begin{cases} |y - x| & \text{if } \phi(x) - \phi(y) \leq \|y - x\|, \\ +\infty & \text{otherwise.} \end{cases}$$

This selection criterion has been applied first to the case of crystalline norms in [2], where in Section 4 one can also find the equivalence of the two minimizations in (4), by a general a variational argument based on Γ-convergence (applied also in the proof of Proposition 7.1 of [3]). The point of this paper is to show how to adapt the disintegration technique from [6, 9] in order to provide an area formula for the disintegration w.r.t. the rays of a plan which is optimal also for the secondary transport problem. Then one can apply the Sudakov-type argument to deduce existence and uniqueness of the optimal transport plan π monotone along rays which solves the secondary problem (4). In particular, this provides a different and simple proof of the existence result in [20]. We try to sketch it after some statements, the proofs are in Section 2.

We obtain more precisely the following. Let $\varepsilon \to 0^+$ and let t_ε be the optimal transport map, non decreasing along rays, between μ and ν for the strictly convex norm

$$c_\varepsilon(x,y) := \|y - x\| + \varepsilon |y - x|.$$

This map satisfies an absolutely continuity push forward estimate (below) that we want to prove in the limit. Restrict the attention for example to any part S of the domain $\{x \cdot \mathrm{e} \leq h^-\}$ where the map t_ε is valued in $\{x \cdot \mathrm{e} \geq h^+\}$. Lemma 2.17 of [17] proves that the maps satisfy the following area estimate:

$$\left(\frac{h^+ - t}{h^+ - s}\right)^{n-1} \mathcal{H}^{n-1}(\sigma_\varepsilon^s S) \leq \mathcal{H}^{n-1}(\sigma_\varepsilon^t S) \leq \left(\frac{t - h^-}{s - h^-}\right)^{n-1} \mathcal{H}^{n-1}(\sigma_\varepsilon^s S), \tag{5}$$

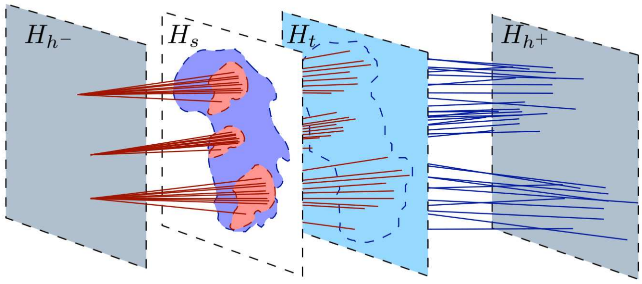


Figure 2: Area estimates of sections. The worst and better cases are obtained when the transport rays (right strip) are rays of cones (left and intermediate strips). Indeed, the proof can be made by cone approximations and a limiting procedure.

where $\sigma^s_\varepsilon S$, $\sigma^t_\varepsilon S$ are the intersections of the segments $[\![x, t_\varepsilon(x)]\!]$, for $x \in S$, with the hyperplanes $H_s = \{x \cdot \mathrm{e} = s\}$, $H_t = \{x \cdot \mathrm{e} = t\}$ and $h^- < s \leq t < h^+$. This estimate means that, moving a transversal section along rays, the area can either increase or decrease at most as if we were moving between H_{h^-} and H_{h^+} along cones with vertices respectively on H_{h^-} or H_{h^+}. See Figure 2.

Theorem 1.1. *The maps t_ε converge in measure to* the c_{s}*-optimal transport map monotone along rays.*

This implies pointwise convergence up to subsequence, and then by the minimality condition one has immediately that $t_\varepsilon - t$ converges to zero in $L^1(\mu)$. By the Γ-convergence argument quoted above ([2]), t and t_ε should moreover satisfy the asymptotic expansion

$$\int \|t(x) - x\|_\varepsilon d\mu = \int \|t_\varepsilon(x) - x\| d\mu + \varepsilon \int |t_\varepsilon(x) - x| d\mu + o(\varepsilon).$$

We sketch now the proof. If μ_ε, ν_ε are finite Radon measures w^*-converging to μ, ν, by the theory of Γ-convergence any w^*-limit π of c_ε-optimal transport plans π_ε in $\Pi(\mu_\varepsilon, \nu_\varepsilon)$ is a c_{s}-optimal transport plan in $\Pi(\mu, \nu)$ (e.g. Th. 4.1 in [3]). The convergence in μ-measure of a sequence of maps t_ε to a map t is equivalent to the w^*-convergence of the plans $(\mathrm{Id}, t_\varepsilon)_\sharp\mu$ to $(\mathrm{Id}, t)_\sharp\mu$: then the theorem follows

- providing the stated regularity of one limit plan π of π_ε (Proposition 1.4);
- observing that it is the unique c_{s}-optimal transport plan monotone along rays (Lemma 1.6).

Remark 1.2. *By uniqueness, the convergence holds also for different approximations, e.g. approximating contemporary ν by finitely many masses ν_ε w^*-converging to ν.*

Before stating these auxiliary results, we remind some standard notations.

Recall 1.3. *By* transport set *associated to a set $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ we refer to the set*

$$\mathcal{T} = \big\{z : \quad z \in [\![x, y]\!],\ (x, y) \in \Gamma\big\},$$

where $[\![x, y]\!]$ denotes the segment from x to y with endpoints, $(\!|x, y|\!)$ without. The transport set associated to a transport plan π is then a transport set associated to some Γ such that $\pi(\Gamma) = 1$. This definition is motivated by the fact that the optimal transport w.r.t. a strictly convex norm cost moves the mass along straight lines, as a consequence of the fact that the triangular inequality is strict when points are not aligned.

A set $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ is c_{s}-monotone *if for all finite number of points $\{(x_i, y_i)\}_{i=1,\dots,M}$ belonging to Γ one has the following inequality:*

$$c_{\mathrm{s}}(x_0, y_0) + \dots + c_{\mathrm{s}}(x_M, y_M) \le c_{\mathrm{s}}(x_0, y_1) + \dots + c_{\mathrm{s}}(x_{M-1}, y_M) + c_{\mathrm{s}}(x_M, y_0).$$

In turn a transport plan π is c_{s}-monotone if there exists a c_{s}-monotone set Γ such that $\pi(\Gamma) = 1$. In this case, secondary rays of π are those (nontrivial) segments $[\![x, y]\!]$ such that $(x, y) \in \Gamma$, where one can assume that

$$(x, y) \in \Gamma,\ [\![z, w]\!] \subset [\![x, y]\!] \quad \Longrightarrow (z, w) \in \Gamma, \tag{6a}$$

$$(x, y), (z, w) \in \Gamma,\ [\![z, w]\!] \cap [\![x, y]\!] = [\![z, y]\!] \quad \Longrightarrow (x, w) \in \Gamma. \tag{6b}$$

Initial/terminal points of secondary rays are usually intended to be the initial/terminal points of secondary rays maximal w.r.t. inclusion. A plan π is monotone (non decreasing) along rays if there exists $\overline{\Gamma}$, $\pi(\overline{\Gamma}) = 1$, such that if $(x, y), (z, w) \in \overline{\Gamma}$ then $[\![z, w]\!] \not\subset (\!|x, y|\!)$, or equivalently $(y - w) \cdot (x - z) \ge 0$ when aligned.

Proposition 1.4. *There exists a c_{s}-optimal transport $\pi \in \Pi(\mu, \nu)$ such that the disintegration of the Lebesgue measure w.r.t. the rays of π has conditional probabilities equivalent to the Hausdorff 1-dimensional measure on the rays. The area estimates* (5) *hold.*

Corollary 1.5. *A c_{s}-optimal transport $\pi \in \Pi(\mu, \nu)$ is induced by a Borel map.*

The corollary is based on the 1-dimensional result: by the disintegration the transport in $\mathbb{R}^n$ reduces to transports on the 1-dimensional rays. The new measures to be transported are the conditional probabilities of μ and ν: then by Proposition 1.4 each conditional probability providing the first marginal on the relative ray is absolutely continuous and therefore the optimal transport problem can be solved by a map (a full proof is e.g. in [17], Th. 3.2). In particular the absolutely continuous disintegration implies that the (measurable) set of initial points is Lebesgue negligible, because its Hausdorff 1-dimensional measure on each ray is 0.

Having Proposition 1.4, at μ-a.e. point there is a unique outgoing secondary ray of π, because by c_{s}-monotonicity there can be more outgoing rays only at initial points and the set of initial points is negligible. Then one can see by considering convex combinations that any other c_{s}-optimal transport plan π' must have that same vector field of secondary rays direction, so that Lemma 1.6 below applies yielding the uniqueness stated in Theorem 1.1.

Lemma 1.6. *If there exists a Borel vector field fixing the direction of the secondary transport ray of any c_{s}-optimal transport plan $\pi' \in \Pi(\eta, \xi)$ at η-a.e. point, there exists at most one c_{s}-optimal plan in $\Pi(\eta, \xi)$ monotone along rays.*

Proofs are provided in Section 2, here we just sketch some ideas. We remark that by [7] the absolute continuity of Proposition 1.4 follows by the simplified area push forward estimate

$$\mu(A) > 0 \qquad \Rightarrow \qquad \mu(A_t) > 0 \text{ for a } \mathcal{L}^1\text{-positive set of times } t,$$

where A_t is the set of points in A translated of length t along any secondary ray they belong to. In this classical setting one can also prove the quantitative full estimate (5), which is clearly stronger.

This push forward area estimate for the secondary rays of the selected transport plan π is the main issue of this paper. It is derived by a compactness argument, and estimates (5) on approximating maps. In particular, it can be obtained as in the literature if one knows that there is just one c_{s}-optimal transference plan monotone along rays (Section 2.2). It is not difficult however to establish that the optimal transport monotone along rays is unique if the direction e.g. of terminal points is fixed for almost every point w.r.t. the target measure (Lemma 2.1, Section 2.1).

We then split our transport plan into partial transports, restrictions of π on suitable regions of $\mathbb{R}^n \times \mathbb{R}^n$. Moreover, we split them into fictional intermediate ones which are easily seen to be the unique c_{s}-optimal transport plans, monotone along rays, among their marginals. More precisely, for a model transport from $\mathbf{B}_{1/2}(0)$ to $\{x \cdot \mathrm{e} \geq 1\}$, the first fictional transport goes from $\mathbf{B}_{1/2}(0)$ to an intermediate section of the transport set, transversal to the rays, with an hyperplane H_λ, while the second from H_λ to $\{x \cdot \mathrm{e} \geq 1\}$. Let η be the target measure of the first one, and the source of the second one (see Figure 3). By standard geometric considerations implied by c_{s}-optimality and then c_{s}-monotonicity, one can deduce that the direction of the first transport is fixed at η-a.e. point. This allows a reduction to the previous cases, yielding the area push forward estimate for these partial transports, restrictions of π.

Having the area push forward estimate everything is done. We recover that the set of initial points is Lebesgue negligible (e.g. Lemma 2.20 in [17], coming from [9]). This implies again the uniqueness of the optimal transport plan monotone along rays (Lemma 1.6 below), and thus full estimates.

In Section 3 we stress some standard consequences of the disintegration result, and of the quantitative estimates. Namely, they provide some regularity of the divergence of rays directions vector field — a kind of Green-Gauss formula holds on special sets — and it allows moreover an explicit expression for the transport density. We give finally an example of the fact that the global optimal Kantorovich potential for the secondary problem with the cost c_{s} does not exist in general, but only on countably many sets which partition μ-all of $\mathbb{R}^n$.

2. Proof

We show the convergence of the optimal maps t_ε, non decreasing along rays, for the transport problem

$$c_\varepsilon(x,y) := \|y-x\| + \varepsilon|y-x|$$

to the optimal map non decreasing along rays for the transport problem

$$\min_{\pi\in\Pi(\mu,\nu)} \int c_{\mathrm{s}}(x,y)\,d\pi(x,y), \quad c_{\mathrm{s}}(x,y) := \begin{cases} |y-x| & \text{if } \phi(x)-\phi(y) \le \|y-x\|, \\ +\infty & \text{otherwise.} \end{cases}$$

Section 2.1 remarks a uniqueness criterion relying on the fact that the direction of the transport is fixed at almost every point w.r.t. the source or target measure, that we state for the case we are considering.

We prove then by stability absolutely continuity area estimates for a limit plan. The estimates are quantitative as in [9], we indeed follow the same basic argument. In Section 2.2 we first prove this estimate in a simpler case, assuming that the support of the w^*-limit of $(\mathrm{Id}, t_\varepsilon)_\sharp\mu$ is c_{s}-cyclically monotone and that μ, ν are concentrated on disjoint balls. Then it is generalized in Section 2.3 by a countable partition satisfying some uniform estimates, and by the uniqueness of these c_{s}-optimal partial transports among their marginals.

2.1. A Uniqueness Remark

Consider two probability measures $\eta, \xi \in \mathcal{P}(\mathbb{R}^n)$. We stress uniqueness of the c_{s}-optimal transport plan (monotone along rays) when the direction of the transport is fixed e.g. at terminal points, by disintegration along rays and uniqueness of the optimal transport map in dimension one — known fact that does not require absolutely continuity assumptions.

LEMMA 2.1. *If there exists a Borel vector field fixing the direction of the secondary transport ray of any c_{s}-optimal transport plan $\pi' \in \Pi(\eta,\xi)$ at η-a.e. point, then there is a unique c_{s}-optimal plan in $\Pi(\eta,\xi)$ monotone along rays.*

Proof. Given any c_{s}-optimal transport plans $\bar\pi, \pi' \in \Pi(\eta,\xi)$, consider a transport set $\mathcal{T}$ of $(\pi' + \bar\pi)/2$.

By c_{s}-monotonicity the secondary rays composing $\mathcal{T}$ can bifurcate only at endpoints. Moreover, by definition of terminal points both π, π' leave them fixed, and thus they coincide there; let us directly assume that the set of terminal points is η-negligible. Since by assumption at η-a.e. initial point of $\mathcal{T}$ there is precisely one secondary transport ray, then the vector field of secondary rays r'_{q} is single valued η-a.e.: secondary rays $\{\mathsf{r}_{\mathsf{q}}\}_{\mathsf{q}}$ partition η-almost all $\mathcal{T}$.

We recall that rays can be parametrized w.l.o.g. by countably many compact subsets of hyperplanes. Let h denote the quotient projection, and $\theta = h_\sharp \eta$ the quotient measure. Consider the disintegration of η, $\bar\pi$, π'

$$\eta = \int \eta_{\mathsf{q}} \theta(\mathsf{q}), \quad \bar\pi = \int \bar\pi_{\mathsf{q}} \theta(\mathsf{q}), \quad \pi' = \int \pi'_{\mathsf{q}} \theta(\mathsf{q})$$

respectively w.r.t. secondary transport rays $\{\mathsf{r}_{\mathsf{q}}\}_{\mathsf{q}}$ and w.r.t. the partition $\{\mathsf{r}_{\mathsf{q}} \times \mathbb{R}^n\}_{\mathsf{q}}$. In particular, we show that $\bar\pi_{\mathsf{q}}$ and π'_{q} have the same second marginal ξ_{q} for θ-a.e. q: indeed for any θ-measurable A and Borel S

$$\begin{aligned}
\xi(h^{-1}(A) \cap S) &\overset{\pi' \in \Pi(\eta,\xi)}{=} \pi'([0,1] \times (h^{-1}(A) \cap S)) \\
&= \int_A \pi'_{\mathsf{q}}([0,1] \times S)\theta(d\mathsf{q}) \\
&= \int_A \xi'_{\mathsf{q}}(S)\theta(d\mathsf{q}),
\end{aligned}$$

and the same holds for $\bar\pi$ with the second marginals $\bar\xi_{\mathsf{q}}$. Therefore, it must be $\bar\xi_{\mathsf{q}} = \xi'_{\mathsf{q}}$ and $\bar\pi_{\mathsf{q}}$, π'_{q} are monotone, 1-dimensional, $|y-x|$-optimal transports in $\Pi(\mu_{\mathsf{q}}, \xi_{\mathsf{q}})$, even if we do not know whether $\xi_{\mathsf{q}} \ll \mathcal{H}^1 \llcorner_{\mathsf{r}_{\mathsf{q}}}$.

By the uniqueness of transport plans in dimension one (see e.g. Prop. 4.5 in [2]), then π and $\bar\pi$ must coincide. $\square$

2.2. Example

Before treating the general case, we consider an example where we assume that the support of the limit plan is c_{s}-monotone. Even if by the theory of Γ-convergence the limit plan π is c_{s}-optimal and then c_{s}-monotone, since the cost c_{s} is just l.s.c. this hypothesis is indeed a restriction: we have $\pi(\Gamma) = 1$ for some c_{s}-monotone set Γ, which can be taken σ-compact but in general not closed. Moreover, if we restrict π to a compact set, then we loose in general the information that it is obtained by a limit.

If we knew the uniqueness of the c_{s}-optimal transport plan π, monotone along rays, then we could instead restrict π to a compact subset of Γ and we could apply to this restriction π' e.g. the statement below. Indeed, this restriction is still the unique c_{s}-optimal transport between its marginals: the transport plans induced by the c_ε-optimal transport maps between the marginals of π' would necessarily w^*-converge to π'.

Focus on an elementary domain. Let $\mu = f\mathcal{L}^n \llcorner_C$ for a compact $C \subset \mathbf{B}_{1/2}(0)$ and $f > 0$ on C, let e be a unit vector. Consider a sequence of continuous transport maps $t_{\varepsilon_j} : C \to \{x \cdot \mathrm{e} \geq 1\}$ which are c_{ε_j}-optimal and such that $(\mathrm{Id}, t_{\varepsilon_j})_\sharp \mu$ is weakly* convergent to a plan π.

We denote by Γ the support of π, extended by (6), and by $\mathcal{T}$ the relative transport set. The flux on secondary rays of π is the multivalued map $x \mapsto \sigma^t(x)$ which moves points along rays defined by

$$\sigma^t(x) = \big\{z : \ (x,z) \in \Gamma, \quad z \cdot \mathrm{e} = x \cdot \mathrm{e} + t\big\}.$$

The domain $\mathrm{Dom}(\sigma^h)$ of σ^h is the set of x such that $\sigma^h(x)$ is nonempty.

Lemma 2.2. *If the support of π is c_{s}-monotone, the transport set $\mathcal{T}$ satisfies the estimate*

$$\left(\frac{1-t}{1-s}\right)^{n-1} \mathcal{H}^{n-1}\big(S\big) \leq \mathcal{H}^{n-1}\big(\sigma^{t-s} S\big) \tag{7}$$

for all compact $S \subset C \cap \{x \cdot \mathrm{e} = s\} \cap \mathrm{Dom}(\sigma^{t-s})$ and $s < t \leq 1$. The symmetric estimate holds similarly.

Corollary 2.3. *If there exists a unique c_{s}-optimal transport plan π, it is induced by a transport map satisfying the absolutely continuous push forward estimates of the kind* (5).

The corollary follows by the elementary restriction argument above.

Proof of Lemma 2.2. Since π is concentrated on the compact set Γ, by the weak convergence for every $\delta > 0$

$$\begin{aligned} 0 &= \pi\big(\big\{(x,y) : \ \mathrm{dist}((x,y),\Gamma) \geq \delta\big\}\big) \\ &\geq \limsup_{\varepsilon_j \to 0} \Big\{ \big((\mathrm{Id}, t_{\varepsilon_j})_\sharp \mu\big)\big(\big\{(x,y) : \ \mathrm{dist}((x,y),\Gamma) \geq \delta\big\}\big) \Big\} \geq 0. \end{aligned}$$

Thus $\mu(\{x : \ \mathrm{dist}((x, t_{\varepsilon_j}(x)),\Gamma) \geq \delta\})$ tends to 0 as $\varepsilon_j \to 0$. Up to a subsequence, one can require then

$$\mu(J_j) < 2^{-j} \qquad \text{with } J_j := \big\{x : \ \mathrm{dist}((x, t_{\varepsilon_j}(x)),\Gamma) > 2^{-j}\big\}.$$

Define the intermediate hyperplanes

$$H_\lambda = \{x:\ x\cdot \mathrm{e} = \lambda\}.$$

Notice that $\mathcal{H}^{n-1}(J_j \cap H_\lambda)$ converges to 0 for $\mathcal{L}^1$-a.e. λ, being $\mathbb{1}_{J_i}$ converging to 0 $\mathcal{L}^n$-a.e.

Up to a translation, we set $s = 0$. Let S be any compact subset of $C \cap \{x\cdot \mathrm{e} = 0\} \cap \mathrm{Dom}(\sigma^t)$. Since secondary transport rays are identified by the compact, c_{s}-monotone set Γ, which is the support of π suitably extended by (6), and since S is also compact, notice then that $\sigma^t(S)$ is compact, too.

Let $h \to \sigma_j^h$ be the analogous flux along secondary rays of t_{ε_j} and set

$$K^j := \sigma_j^t(S \setminus J_j) \quad \subset \quad \{x\cdot \mathrm{e} = t\}.$$

Being compact, K^j converges in the Hausdorff distance, up to subsequence, to a compact set K. Moreover, since by construction $d((x, t_{\varepsilon_j}(x)), \Gamma) \leq 2^{-j}$ out of J_j, we have that $K \subset \overline{\sigma^t(S)} = \sigma^t(S)$. By the u.s.c. of the Hausdorff measure and the regularity of the approximating vector field we conclude

$$\begin{aligned}\mathcal{H}^{n-1}(\sigma^t(S)) \geq \mathcal{H}^{n-1}(K) &\geq \limsup_j \mathcal{H}^{n-1}(K^j)\\ &\geq \lim_j \big\{(1-t)^{n-1}\mathcal{H}^{n-1}(S\setminus J_j)\big\} = (1-t)^{n-1}\mathcal{H}^{n-1}\big(S\big),\end{aligned}$$

where the last equality holds if $\mathcal{H}^{n-1}(J_j \cap H_\lambda)$ goes to 0. The thesis holds as well also for the remaining ($\mathcal{L}^1$-negligible) values of s by the lower semicontinuity of $\mathcal{H}^{n-1}$, being for λ decreasing to s

$$\mathcal{H}^{n-1}(\mathbf{B}_\delta(S)\cap H_s) \leq \liminf_{\lambda\downarrow s} \mathcal{H}^{n-1}(\mathbf{B}_\delta(\sigma^{\lambda-s}S)\cap H_\lambda). \qquad \square$$

2.3. Proof of Proposition 1.4

We disintegrate here the Lebesgue measure on the transport set $\mathcal{T}$, associated to any w^*-limit π of $(\mathrm{Id}, t_{\varepsilon_j})_\sharp\mu$, w.r.t. the partition into secondary rays: we show by a quantitative area push forward estimate that the conditional probabilities are absolutely continuous. We basically reduce to the case of the example in Lemma 2.2.

The idea is the following. In the model case of a transport $\mathbf{B}_{1/2}(0) \to \{x\cdot \mathrm{e} \geq 1\}$, we see the optimal transport plan we selected as a composition of two other optimal transport plans: $\mathbf{B}_{1/2}(0) \to H_\lambda$ and $H_\lambda \to \{x\cdot \mathrm{e} \geq 1\}$, where H_λ is an intermediate section transversal to the rays. The two intermediate transports should still be c_{s}-optimal, and moreover they basically share the same secondary rays as their composite plan. The terminal points of secondary rays of the first, coinciding with the initial ones of the second, should then be

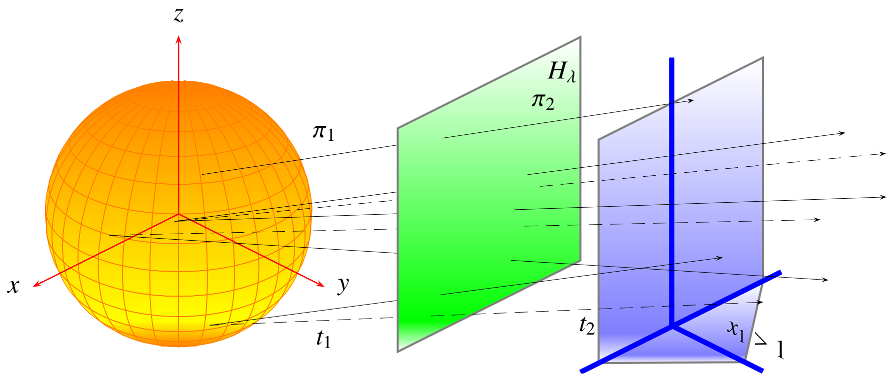


Figure 3: Decomposition of the transport.

fixed — because of c_{s}-monotonicity applied to the composite plans. Then the intermediate transports we selected are the unique optimal transports between their marginals by Lemma 1.6. In particular, Lemma 2.2 applies for these transports yielding the area estimate, which holds both for the intermediate and composite transports.

Since the property of the proposition is local, up to a countable partition and similarity transformations we are allowed to decompose π into countably many restrictions of it, that we place on a model set. Notice moreover that, up to the present purpose, one does not need to consider the restriction of π to the diagonal, i.e. the fixed points, because also removing them the secondary rays of the transport set remain the same. They matter only in order to determine later the transport map solving the Monge problem with the given marginal.

Renewing the notations for the marginals of these partial plans, we assume

$$\mu \ll \mathcal{L}^n \llcorner_{\mathbf{B}_{1/2}(0)} \qquad \text{and} \qquad \nu(\{x \cdot \mathrm{e} \geq 1\}) = 1.$$

We now denote by t_ε the c_ε-optimal maps monotone along ray from μ to ν. If, by Lusin theorem, we also assume that the maps t_ε are continuous by restricting them to suitable compact sets, then we are in a setting where the second marginal is in general different from ν, w^*-converging to it:

$$t_\varepsilon : \mathbf{B}_{1/2}(0) \to \{x \cdot \mathrm{e} \geq 1\} \qquad \text{and} \qquad \nu_\varepsilon := (t_\varepsilon)_\sharp \mu.$$

The w^*-limit π of a sequence $(\mathrm{Id}, t_\varepsilon)_\sharp \mu$, after showing the uniqueness, will turn out to be precisely one of the restrictions originally considered.

View each map t_ε as the composition of two c_ε-optimal transports (Figure 3): for $\lambda \in (1/2, 1)$

$$t^1_\varepsilon(x) := x + \lambda[(\mathrm{e} - x) \cdot d_\varepsilon(x)]d_\varepsilon(x), \qquad t^2_\varepsilon(x) := t_\varepsilon((t^1_\varepsilon)^{-1}(x)),$$

$$d_\varepsilon(x) := \frac{t_\varepsilon(x) - x}{|t_\varepsilon(x) - x|}.$$

Let $\xi_\varepsilon := (t^1_\varepsilon)_\sharp \mu$ be the intermediate measure on $H_\lambda := \{x : \ x \cdot \mathrm{e} = \lambda\}$, which is the source of t^2_ε and the target of t^1_ε. Notice that t^1_ε is injective, by the c_ε-monotonicity, so that t^2_ε is well defined ξ_ε-a.e.

By compactness, the transport plans π^2_ε, π^1_ε associated to t^1_ε, t^2_ε w^*-converge, up to a subsequence, to plans $\pi^1 \in \Pi(\mu, \xi)$, $\pi^2 \in \Pi(\xi, \nu)$ — where ξ, ν are the w^*-limit of ξ_ε, ν_ε. By the theory of Γ-convergence (see e.g. Th. 4.1 in [3]) π^1, π^2 are c_s-optimal transport plans. Moreover, since for all $\hat\pi \in \Pi(\mu + \xi_\varepsilon, \nu_\varepsilon + \xi_\varepsilon)$

$$\int c_\varepsilon(x, t^1_\varepsilon(x))\mu(dx) + \int c_\varepsilon(z, t^2_\varepsilon(z))\xi(dz) = \int c_\varepsilon(x, t_\varepsilon(x))\mu(dx) \leq \int c_\varepsilon \hat\pi,$$

by the c_ε-optimality of $(\mathrm{Id}, t^1_\varepsilon)_\sharp \mu + (\mathrm{Id}, t^2_\varepsilon)_\sharp \xi_\varepsilon$ also $\pi^1 + \pi^2$ *is c_s-optimal.*

Since (maximal) secondary rays of $\pi^1 + \pi^2$ go from $\mathbf{B}_{1/2}(0)$ to $\{x \cdot \mathrm{e} \geq 1\}$, *the direction of the transport is unique at ξ-a.e. point.* The measurability of this vector field follows from the fact that there is a representative with a σ-compact graph, because $\pi^1 + \pi^2$ is concentrated on a σ-compact set. Observing that $\pi^1 \ll \pi^1 + \pi^2$ and that π^1 is monotone along rays, then Lemma 1.6 states that π^1 is the only c_s-optimal transport plan from μ to ξ monotone along rays.

The uniqueness of π^1 yields the basic push forward estimate (7) for π^1 by Corollary 2.3. However, this estimate coincides with the basic push forward estimate also for the c_s-optimal transports $\pi^1 + \pi^2 \in \Pi(\mu + \xi, \nu + \xi)$ and for

$$\Pi(\mu, \nu) \ni \pi(dx, dy) := \int \pi^2_z(dy)\pi^1(dx, dz), \text{ where } \pi^2(dz, dy) = \int \pi^2_z(dy)\xi(dz),$$

which share the same transport set $\mathcal{T}$.

This yields a one-sided estimate, but it is enough in order to deduce by a density argument that the set of initial points of π is negligible (precisely Lemma 2.20 in [17], proved before in [9]). Indeed, if we could take a point $\bar x$ of density one for the set of initial points $\mathcal{E}$, then we would reach an absurd: on one hand moving them along the ray directions, close to e, they are no more initial points and therefore they do not belong to the set $\mathcal{E}$; on the other hand the one-sided estimate implies that $\bar x$ is a point of positive density for these translated points.

As remarked in Section 2.1, then uniqueness hold and one finds by the limiting procedure the full estimates giving the statements of the proposition.

3. Some Remarks

The first corollary of the previous computations is the following disintegration.

THEOREM 3.1. *The family of secondary transport rays* $\{\mathsf{r}_\mathsf{q}\}_{\mathsf{q}\in\mathsf{Q}}$ *can be parameterized by a Borel subset* Q *of countably many hyperplanes. The transport set* $\mathcal{T}=\cup_{\mathsf{q}\in\mathsf{Q}}\mathsf{r}_\mathsf{q}$ *is Borel and there exists a Borel function* γ *such that the following disintegration of* $\mathcal{L}^n \llcorner_{\mathcal{T}}$ *holds:* $\forall\varphi$ *either integrable or positive*

$$\int_{\mathcal{T}}\varphi(x)d\mathcal{L}^n(x)=\int_{\mathsf{Q}}\left\{\int_{\mathsf{r}_\mathsf{q}}\varphi(s)\gamma(s)d\mathcal{H}^1(s)\right\}d\mathcal{H}^{n-1}(\mathsf{q})$$

The set of endpoints of rays is Lebesgue negligible.

Denote by d the unit vector field of secondary rays directions, defined $\mathcal{L}^n$-a.e. on $\mathcal{T}$. The quantitative estimates, as in the previous works in $\mathbb{R}^n$, imply a further regularity of the density γ. We omit the proof and precise formulas, analogous to the one of Proposition 4.17 in [16].

LEMMA 3.2. *For* $\mathcal{L}^n$*-a.e.* x *the real function*

$$\lambda\mapsto\gamma(x+\lambda\mathsf{d}(x))$$

is locally Lipschitz, with locally finite total variation, for x, $x+\lambda\mathsf{d}(x)$ *belonging to a same ray.*

Moreover, as e.g. in [16] the function

$$\frac{\partial_\lambda\gamma(x+\lambda\mathsf{d}(x))}{\gamma(x+\lambda\mathsf{d}(x))}=:\big(\operatorname{div}\mathsf{d}\ \big)_{\text{a.c.}}(x)$$

is of particular interest, as we explain below motivating the abuse of notation.

3.1. Divergence of the Rays Directions Vector Field

Consider a compact subset $\mathcal{Z}$ of the transport set $\mathcal{T}$ made of secondary transport rays which intersect an hyperplane H at points which are in the relative interior of the rays. Then the divergence of the vector field $\mathsf{d}\mathbb{1}_{\mathcal{Z}}$ is a Radon measure.

LEMMA 3.3. *There exist nonnegative measures* η^+,η^-*, concentrated respectively on initial and terminal points of secondary rays of* $\mathcal{Z}$*, such that*

$$\operatorname{div}(\mathsf{d}\mathbb{1}_{\mathcal{Z}})=\big(\operatorname{div}\mathsf{d}\ \big)_{\text{a.c.}}(x)\mathcal{L}^n\llcorner_{\mathcal{Z}}-\eta^++\eta^-.$$

If the initial points are on a same hyperplane H^- *and the terminal points on a same hyperplane* H^+ *orthogonal to a unit direction* e*, then the measures* $\mu^\pm$ *are just* $|\mathsf{d}\cdot\mathrm{e}|\mathcal{H}^{n-1}$ *on the relative hyperplane.*

The proof relies on the disintegration theorem for reducing the integrals $\int_{\mathcal{Z}} \nabla\varphi \cdot \mathrm{d}\mathcal{L}^n$ on the rays, where φ is a test function. The factor γ appears in the area formula, and *on each ray* by the estimates providing BV regularity one can integrate by parts (see the proof of Lemma 2.30 in [17]).

It follows then that the distributional divergence of d is a series of measures. We remark however that in general the divergence of d is just a distribution, and it may fail to be a measure (see e.g. Examples 4.2, 4.3 in [17]). As well, the function $\big(\operatorname{div}\mathsf{d}\big)_{\text{a.c.}}$ could fail to be locally integrable.

3.2. Transport Density

We now stress another known consequence of the disintegration theorem: one can write the expression of the transport density, vanishing approaching initial points along secondary transport rays, relative to optimal secondary transport plans in terms of the conditional measures $\mu_{\mathtt{q}}$, $\nu_{\mathtt{q}}$, $\mathtt{q} \in \mathtt{Q}$ of μ, ν for the ray equivalence relation. In particular, one can see its absolute continuity. It does not vanish approaching terminal points — see Example 3.5 below taken from [24]. We omit the verification, since it is quite standard (see e.g. Section 8 in [9]).

Let f, the Radon-Nycodim derivative of μ w.r.t. $\mathcal{L}^n$, and γ, introduced in the disintegration, be Borel functions from $\mathbb{R}^n$ to $\mathbb{R}$ such that

$$\mu\llcorner\mathcal{T} = \int_{\mathtt{Q}} \mu_{\mathtt{q}}\, d\mathcal{H}^{n-1}(\mathtt{q}) = \int_{\mathtt{Q}} (f\gamma\mathcal{H}^1\llcorner_{r_{\mathtt{q}}})\, d\mathcal{H}^{n-1}(\mathtt{q}) \qquad \nu\llcorner\mathcal{T} = \int_{\mathtt{Q}} \nu_{\mathtt{q}}\, d\mathcal{H}^{n-1}(\mathtt{q}).$$

Let $\mathtt{q} : \mathcal{T} \to \mathtt{Q}$ be the Borel multivalued quotient projection. Set $d = 0$ where $\mathcal{D}$ is multivalued.

LEMMA 3.4. *A solution $\rho \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^n)$ to the transport equation*

$$\operatorname{div}(\rho\mathsf{d}) = \mu - \nu$$

is given by

$$\rho(x) = \frac{(\mu_{\mathtt{q}(x)} - \nu_{\mathtt{q}(x)})((\!|a(x), x|\!))}{\gamma(x)} \mathcal{L}^n(x)\llcorner\mathcal{T} = \left(\frac{\mathbb{1}_{\mathcal{T}}(x)}{\gamma(x)} \int_{(\!|\tilde{t}^{-1}(x),x|\!)} f\gamma\, d\mathcal{H}^1\right)\mathcal{L}^n(x). \tag{8}$$

EXAMPLE 3.5 (Taken from [24]). *Consider in $\mathbb{R}^2$ the measures $\mu = 2\mathcal{L}^2\llcorner_{\mathbf{B}_1}$ and $\nu = \frac{1}{2|x|^{3/2}}\mathcal{L}^2\mathbf{B}_1$, where $|\cdot|$ here denotes the Euclidean norm. A Kantorovich potential is provided by $|x|$. The transport density is $\rho = (|x|^{-\frac{1}{2}} - |x|)\mathcal{L}^2\llcorner_{\mathbf{B}_1}$. While vanishing towards $\partial\mathbf{B}_1$, the density of ρ blows up towards the origin. Concentrating ν at the origin, the density would be instead $\rho = -|x|^2\llcorner_{\mathbf{B}_1}$.*

3.3. Example of no Global Secondary Potential

We show here that in general there exists no function ϕ_{s} which satisfies on the whole transport set relative to two measures μ, ν

$$\phi_{\mathrm{s}}(x) - \phi_{\mathrm{s}}(y) = c_{\mathrm{s}}(x, y) \qquad \text{for } x, y \text{ on a same secondary ray} \tag{9a}$$

$$\phi_{\mathrm{s}}(x) - \phi_{\mathrm{s}}(y) \leq c_{\mathrm{s}}(x, y) \qquad \forall (x, y) \in \mathcal{T} \times \mathcal{T}. \tag{9b}$$

It would otherwise provide a global Kantorovich potential for the secondary transport problem, which exists only up to a countable partition of the domain. The secondary cost function was defined by

$$c_{\mathrm{s}}(x, y) := \begin{cases} |y - x| & \text{if } \phi(x) - \phi(y) = \|y - x\|, \\ +\infty & \text{otherwise.} \end{cases}$$

Consider in $\mathbb{R}^2$ the norm $\|x\| = |P_1 x| + |P_2 x|$, where P_1, P_2 are the projections on the first and second component, and let $|\cdot|$ denote the Euclidean norm. We show for simplicity of notations a transport problem with atomic marginals, the example can then be adapted spreading the mass as in the pictures of Figure 4. Consider the transport among the measures

$$\mu = \sum_{i=1}^{\infty} \sum_{j=-1}^{4/h_i} (h_i/12)^2 \delta_{w_{ij}} + (1/24)^2 \delta_{w_\infty} \quad \nu = \sum_{i=1}^{\infty} \sum_{j=-1}^{4/h_i} (h_i/12)^2 \delta_{z_{ij}} + (1/24)^2 \delta_{z_\infty}$$

where $h_i = 2^{-i-1}$, $w_\infty = (-1.5, 0)$, $z_\infty = (-1.5, 1)$ and

$$\begin{cases} w_{1,-1} = (0,\, 0) \\ w_{i,-1} = \big(-\sum_{k=1}^{i,-1} 2h_k,\, 0 \big) \\ w_{ij} = w_{i,-1} + (-h_i,\, jh_i/4) \end{cases} \qquad \begin{cases} z_{i,-1} = w_{i,-1} + (0, 1) & i \in \mathbb{N} \\ z_{ij} = w_{ij} + (h_i/2,\, 0) & j = 0, \dots, 4/h_i. \end{cases}$$

Let π be the transport plan induced by the map t which translates each w_{ik} to z_{ik}. It is an optimal one for the primary problem: one can see it for example by duality, noticing that the function

$$\phi(x) = \|x - z_{1,-1}\|$$

is a Kantorovich potential. Moreover, one can immediately verify that it is also c_{s}-optimal. In this case, one can take

$$c_{\mathrm{s}}(x, y) := \begin{cases} |y - x| & \text{if } \{P_1 y \geq P_1 x,\ P_2 y \geq P_2 x\}, \\ +\infty & \text{otherwise.} \end{cases}$$

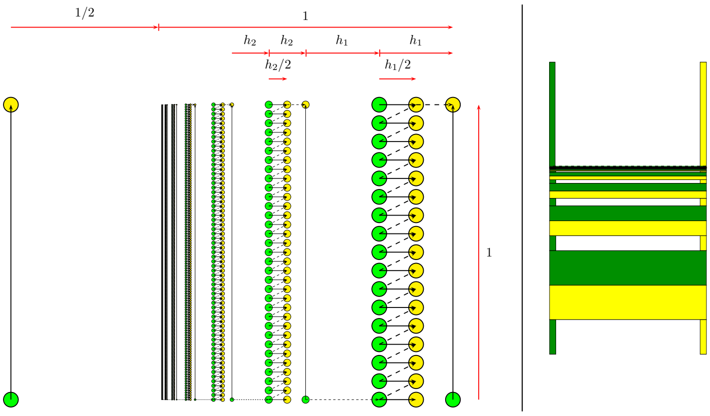


Figure 4: Non existence of a global secondary potential. LHS: back on the path along the arrows the secondary potential must go to $-\infty$. RHS (rotated): the mass is spread so that the primary potential is unique, up to constants.

However, no c_{s}-monotone carriage Γ of π is contained in the c_{s}-subdifferential of a c_{s}-monotone function ϕ_{s}, which by definition would satisfy (9). Indeed, suppose the contrary. Then, considering the path in the figure and applying repeatedly the maximal growth equality (9a) (full line) and the Lipschitz inequality (9b) (dashed line), one finds

$$\begin{aligned}\phi_{\mathrm{s}}(w_{(i+1),-1}) &\leq \phi_{\mathrm{s}}(w_{i,-1}), -1 + \frac{h_i}{2} + \frac{h_i}{2}\left(\frac{\sqrt{5}}{2} - 1\right) \cdot \frac{4}{h_i} + \frac{3h_i}{2} \\ &= \phi_{\mathrm{s}}(w_{i,-1}) + \sqrt{5} - 3 + 2h_i.\end{aligned}$$

For every potential ϕ_{s} finite on $w_{1,-1}$, we find therefore that $\phi_{\mathrm{s}}(w_{i,-1}) \to -\infty$ for $i \to \infty$, as well as every other $\phi_{\mathrm{s}}(w_{ij})$. This implies that ϕ_{s} must be $-\infty$ on w_∞: for all i, j

$$\phi_{\mathrm{s}}(w_\infty) \leq \phi_{\mathrm{s}}(w_{ij}) + \|w_{ij} - w_\infty\|,$$

which implies $\phi_{\mathrm{s}}(w_\infty) = -\infty$.

REMARK 3.6. *One could think that the problem is that the primary potential we have chosen is not the right one. However, this is not the case. A completely similar behavior happens spreading the mass as in the second picture of Figure 4 (rotated of $-\pi/2$), but there the primary potential is unique, up to constants.*

Acknowledgements. The author wishes to thank in particular professor Stefano Bianchini for several valuable discussions and the referee for his constructive remarks.

References

[1] L. Ambrosio, *Lecture notes on optimal transport problems*, Lecture Notes in Mathematics, volume 1812, Springer, Berlin (2003), pp. 1–52.

[2] L. Ambrosio, B. Kircheim and A. Pratelli, *Existence of optimal transport maps for crystalline norms*, Duke Math. J. **125** (2004), 207–241.

[3] L. Ambrosio and A. Pratelli, *Existence and stability results in the* L^1 *theory of optimal transportation*, Lecture Notes in Mathematics, volume 1813, Springer, Berlin (2003), pp. 123–160.

[4] P. Appell, *Mémoire sur les déblais et les remblais des systèmes continus ou discontinus*, Mémoires preéséntes par divers Savants à l'Académie des Sciences de l'Institut de France **29** (1887), 1–208.

[5] M. Beckmann, *A continuous model of transportation*, Econometrica **20** (1952), 643–660.

[6] S. Bianchini, *On the Euler-Lagrange equation for a variational problem*, Discrete Contin. Dyn. Syst. **17** (2007), 449–480.

[7] S. Bianchini and F. Cavalletti, *The Monge problem for distance cost in geodesic spaces*, arXiv:1103.2796.

[8] S. Bianchini and S. Daneri, to appear.

[9] S. Bianchini and M. Gloyer, *On the Euler Lagrange equation for a variational problem: the general case II*, Math. Z. **265** (2010), 889–923.

[10] G. Bouchitté and G. Buttazzo, *Characterization of optimal shapes and masses through Monge-Kantorovich equation*, J. Eur. Math. Soc. **3** (2001), 139–168.

[11] G. Bouchitté, G. Buttazzo and P. Seppecher, *Shape optimization solutions via Monge-Kantorovich equation*, C. R. Acad. Sci. Paris Sér. I Math. **324** (1997), 1185–1191.

[12] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), 805–808.

[13] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), 375–417.

[14] L. Caffarelli, M. Feldman and R. McCann, *Constructing optimal mass for Monge's transport problem as a limit of strictly convex costs*, J. Amer. Math. Soc. (2002), 1–26.

[15] L.A. Caffarelli, *Boundary regularity of maps with convex potentials. II*, Ann. Math. **144** (1996), 453–496.

[16] L. Caravenna and S. Daneri, *The disintegration of the Lebesgue measure on the faces of a convex function.*, J. Funct. Anal. **258** (2010), 3604–3661.

[17] L. Caravenna, *A proof of Sudakov theorem with strictly convex norms*, Math. Z. **268** (2011), 371–407.

[18] F. Cavalletti, *A strategy for non-strictly convex distance transport cost and the obstacle problem*, arXiv:1103.2797.

[19] T. Champion and L. De Pascale, *The Monge problem for strictly convex norms in* $\mathbb{R}^d$, J. Eur. Math. Soc. **6** (2010), 1355-1369.

[20] T. Champion and L. De Pascale, *The Monge problem in* $\mathbb{R}^d$, Duke Math. J. **157** (2011), 551–572.

[21] C. Dupin, *Applications de la géométrie et de la mécanique*, Bachelier, Paris (1822).

[22] L.C. Evans and W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Current Developments in Mathematics (1997), 65–126.

[23] M. Feldman and R. McCann, *Monge's transport problem on a Riemannian manifold*, Trans. Amer. Math. Soc. (2002), 1667–1697.

[24] M. Feldman and R.J. McCann, *Uniqueness and transport density in Monge's mass transportation problem*, Calc. Var. Partial Differential Equations **15** (2002), 81–113.

[25] W. Gangbo, *An elementary proof of the polar factorization of vector-valued functions*, Arch. Rational Mech. Anal. **128** (1994), 381–399.

[26] L.V. Kantorovich, *On the transfer of masses*, Docl. Akad. Nauk. SSSR **37** (1942), 227–229.

[27] L.V. Kantorovich, *On a problem of Monge*, Uskpekhi Mat. Nauk. **3** (1948), 225–226.

[28] M. Knott and C. Smith, *On the optimal mapping of distributions*, J. Optim. Theory Appl. **43** (1984), , 39–49.

[29] D.G. Larman, *A compact set of disjoint line segments in $\mathbb{R}^3$ whose end set has positive measure.*, Mathematika **18** (1971), 112–125.

[30] R.J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J. **80** (1995), 309–323.

[31] G. Monge, *Mémoire sur la theorie des déblais et des remblais*, Histoire de l'Acad. de Sciences de Paris (1781), 666–704.

[32] L. Rüschendorf and S.T. Rachev, *A characterization of random variables with minimum L^2-distance*, J. Multivariate Anal. **32** (1990), 48–54.

[33] F. Santambrogio, *Absolute continuity and summability of transport densities: simpler proofs and new estimates*, Calc. Var. Partial Differential Equations **36** (2008), 343–354.

[34] V.N. Sudakov, *Geometric problems in the theory of infinite-dimensional probability distributions*, Proc. Steklov Inst. Math. **2** (1979), 1–178 (in English); Trudy Mat. Inst. Steklov **141** (1976), 3–191 (in Russian).

[35] N.S. Trudinger and X.J. Wang, *On the Monge mass transfer problem*, Calc. Var. Partial Differential Equations **13** (2001), 19–31.

Author's address:

Laura Caravenna
CRM De Giorgi, Collegio Puteano
Scuola Normale Superiore
Piazza dei Cavalieri 3, 56100 Pisa, Italy
E-mail: `laura.caravenna@sns.it`

Received December 14, 2010
Revised March 21, 2011

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 53–60

Weakly ωb-Continuous Functions[1]

Jamal M. Mustafa

Abstract. *In this paper we introduce a new class of functions called weakly ωb-continuous functions and investigate several properties and characterizations. Connections with other existing concepts, such as ωb-continuous and weakly b-continuous functions, are also discussed.*

Keywords: b-Open Sets, ωb-Open Sets, Weakly b-Continuous Functions, Weakly ωb-Continuous Functions.
MS Classification 2010: 54C05, 54C08

1. Introduction

The notion of $b-open$ sets in topological spaces was introduced in 1996 by Andrijevic [1]. This type of sets discussed by El-Atik [2] under the name of $\gamma-open$ sets. In 2008, Noiri, Al-Omari and Noorani [4] introduced the notions of $\omega b-open$ sets and ωb-continuous functions. We continue to introduce and study properties and characterizations of weakly ωb-continuous functions.

Let A be a subset of a space (X, τ). The closure (resp. interior) of A will be denoted by $Cl(A)$ (resp. $Int(A)$).

A subset A of a space (X, τ) is called $b-open$ [1] if $A \subseteq Cl(Int(A)) \cup Int(Cl(A))$. The complement of a $b-open$ set is called a $b-closed$ set. The union of all $b-open$ sets contained in A is called the $b-interior$ of A, denoted by $bInt(A)$ and the intersection of all $b-closed$ sets containing A is called the $b-closure$ of A, denoted by $bCl(A)$. The family of all $b-open$ (resp. $b-closed$) sets in (X, τ) is denoted by $BO(X)$ (resp. $BC(X)$).

Definition 1.1. *A subset A of a space X is said to be $\omega b-open$ [4] if for every $x \in A$, there exists a $b-open$ subset $U_x \subseteq X$ containing x such that $U_x - A$ is countable.*

The complement of an $\omega b-open$ set is said to be $\omega b-closed$ [4]. The intersection of all $\omega b-closed$ sets of X containing A is called the $\omega b-closure$ of A and is denoted by $\omega bCl(A)$. The union of all $\omega b-open$ sets of X contained in A is called the $\omega b-interior$ of A and is denoted by $\omega bInt(A)$.

[1] Jamal M. Mustafa did this research during the sabbatical leave from Department of Mathematics, Al al-Bayt University, Mafraq, Jordan.

LEMMA 1.2 ([4]). *For a subset of a topological space, b-opennes implies $\omega b-$ openness.*

LEMMA 1.3 ([4]). *The intersection of an $\omega b-$ open set with an open set is $\omega b-$ open.*

LEMMA 1.4 ([4]). *The union of any family of $\omega b-$ open sets is $\omega b-$ open.*

2. Weakly ωb-Continuous Functions

DEFINITION 2.1. *A function $f:(X,\tau)\rightarrow(Y,\rho)$ is said to be:*

(a) *ωb-continuous [4] if for each $x\in X$ and each open set V in Y containing $f(x)$, there exists an $\omega b-$ open set U in X containing x such that $f(U)\subseteq V$.*

(b) *weakly b-continuous [7] if for each $x\in X$ and each open set V in Y containing $f(x)$, there exists a b–open set U in X containing x such that $f(U)\subseteq Cl(V)$.*

DEFINITION 2.2. *A function $f:(X,\tau)\rightarrow(Y,\rho)$ is said to be weakly ωb-continuous if for each $x\in X$ and each open set V in Y containing $f(x)$, there exists an $\omega b-$ open set U in X containing x such that $f(U)\subseteq Cl(V)$.*

REMARK 2.3. *Every ωb-continuous function is weakly ωb-continuous, but the converse is not true in general as the following example shows.*

EXAMPLE 2.4. *Let $X=R$ with the usual topology τ and $Y=\{a,b\}$ with $\rho=\{\phi,Y,\{a\}\}$. Define a function $f:(X,\tau)\rightarrow(Y,\rho)$ by $f(x)=a$ if $x\in Q$ and $f(x)=b$ if $x\in R-Q$. Then f is weakly ωb-continuous but not ωb-continuous.*

REMARK 2.5. *Since every $b-$ open set is $\omega b-$ open then every weakly b-continuous function is weakly ωb-continuous but the converse is not true in general as the following example shows.*

EXAMPLE 2.6. *Let $X=Y=\{a,b,c\}$, $\tau=\{\phi,X,\{c\},\{a,b\},\{a,b,c\}\}$ and $\rho=\{\phi,Y,\{a,b\},\{c,d\}\}$. Define a function $f:(X,\tau)\rightarrow(Y,\rho)$ by $f(a)=a$, $f(b)=d$, $f(c)=c$ and $f(d)=b$. Then f is weakly ωb-continuous but not weakly b-continuous.*

THEOREM 2.7. *A function $f:(X,\tau)\rightarrow(Y,\rho)$ is weakly ωb-continuous if and only if for every open set V in Y, $f^{-1}(V)\subseteq\omega bInt[f^{-1}(Cl(V))]$.*

Proof.

$\Rightarrow$) Let $V\in\rho$ and $x\in f^{-1}(V)$. Then there exists an $\omega b-open$ set U in X such that $x\in U$ and $f(U)\subseteq Cl(V)$. Therefore, we have $x\in U\subseteq f^{-1}(Cl(V))$ and hence $x\in\omega bInt[f^{-1}(Cl(V))]$ which means that $f^{-1}(V)\subseteq\omega bInt[f^{-1}(Cl(V))]$.

$\Leftarrow$) Let $x \in X$ and $V \in \rho$ with $f(x) \in V$. Then $x \in f^{-1}(V) \subseteq \omega bInt[f^{-1}(Cl(V))]$. Let $U = \omega bInt[f^{-1}(Cl(V))]$. Then U is $\omega b - open$ and $f(U) \subseteq Cl(V)$. $\square$

THEOREM 2.8. *Let $f : (X, \tau) \to (Y, \rho)$ be a weakly ωb-continuous function. If V is a clopen subset of Y, then $f^{-1}(V)$ is $\omega b - open$ and $\omega b - closed$ in X.*

Proof. Let $x \in X$ and V be a clopen subset of Y such that $f(x) \in V$. Then there exists an $\omega b - open$ set U in X containing x such that $f(U) \subseteq Cl(V)$. Hence $x \in U$ and $f(U) \subseteq V$ and so $x \in U \subseteq f^{-1}(V)$. This shows that $f^{-1}(V)$ is $\omega b - open$ in X. Since $Y - V$ is a clopen set in Y, so $f^{-1}(Y - V)$ is $\omega b - open$ in X. But $f^{-1}(Y - V) = X - f^{-1}(V)$. Therefore $f^{-1}(V)$ is $\omega b - closed$ in X. Hence $f^{-1}(V)$ is $\omega b - open$ and $\omega b - closed$ in X. $\square$

THEOREM 2.9. *A function $f : (X, \tau) \to (Y, \rho)$ is weakly ωb-continuous if and only if for every closed set C in Y, $\omega bCl[f^{-1}(Int(C))] \subseteq f^{-1}(C)$.*

Proof.

$\Rightarrow$) Let C be a closed set in Y. Then $Y - C$ is an open set in Y so by Theorem 2.8 $f^{-1}(Y - C) \subseteq \omega bInt[f^{-1}(Cl(Y - C))] = \omega bInt[f^{-1}(Y - Int(C))] = X - \omega bCl[f^{-1}(Int(C))]$. Thus $\omega bCl[f^{-1}(Int(C))] \subseteq f^{-1}(C)$.

$\Leftarrow$) Let $x \in X$ and $V \in \rho$ with $f(x) \in V$. So $Y - V$ is a closed set in Y. So by assumption $\omega bCl[f^{-1}(Int(Y - V))] \subseteq f^{-1}(Y - V)$. Thus $x \notin \omega bCl[f^{-1}(Int(Y - V))]$. Hence there exists an $\omega b - open$ set U in X such that $x \in U$ and $U \cap f^{-1}(Int(Y - V)) = \phi$ which implies that $f(U) \cap Int(Y - V) = \phi$. Then $f(U) \subseteq Y - Int(Y - V)$, so $f(U) \subseteq Cl(V)$, which means that f is weakly ωb-continuous. $\square$

THEOREM 2.10. *Let $f : (X, \tau) \to (Y, \rho)$ be a surjection function such that $f(U)$ is $\omega b - open$ in Y for any $\omega b - open$ set U in X and let $g : (Y, \rho) \to (Z, \sigma)$ be any function. If gof is weakly ωb-continuous then g is weakly ωb-continuous.*

Proof. Let $y \in Y$. Since f is surjection, there exists $x \in X$ such that $f(x) = y$. Let $V \in \sigma$ with $g(y) \in V$, so $(gof)(x) \in V$. Since gof is weakly ωb-continuous there exists an $\omega b - open$ set U in X containing x such that $(gof)(U) \subseteq Cl(V)$. By assumption $H = f(U)$ is an $\omega b - open$ set in Y and contains $f(x) = y$. Thus $g(H) \subseteq Cl(V)$. Hence g is weakly ωb-continuous. $\square$

DEFINITION 2.11. *A function $f : (X, \tau) \to (Y, \rho)$ is called ωb-irresolute if $f^{-1}(V)$ is $\omega b - open$ in (X, τ) for every $\omega b - open$ set V in (Y, ρ).*

THEOREM 2.12. *If $f : (X, \tau) \to (Y, \rho)$ is ωb-irresolute and $g : (Y, \rho) \to (Z, \sigma)$ is weakly ωb-continuous then $gof : (X, \tau) \to (Z, \sigma)$ is weakly ωb-continuous.*

Proof. Let $x \in X$ and $V \in \sigma$ such that $(gof)(x) = g(f(x)) \in V$. Let $y = f(x)$. Since g is weakly ωb-continuous. So there exists an $\omega b - open$ set W in Y such that $y \in W$ and $g(W) \subseteq Cl(V)$. Let $U = f^{-1}(W)$. Then U is an $\omega b - open$ set in X as f is ωb-irresolute. Now $(gof)(U) = g(f(f^{-1}(W))) \subseteq g(W)$. Then $x \in U$ and $(gof)(U) \subseteq Cl(V)$. Hence gof is weakly ωb-continuous. □

THEOREM 2.13. *If $f : (X, \tau) \to (Y, \rho)$ is weakly ωb-continuous and $g : (Y, \rho) \to (Z, \sigma)$ is continuous then $gof : (X, \tau) \to (Z, \sigma)$ is weakly ωb-continuous.*

Proof. Let $x \in X$ and W be an open set in Z containing $(gof)(x) = g(f(x))$. Then $g^{-1}(W)$ is an open set in Y containing $f(x)$. So there exists an $\omega b - open$ set U in X containing x such that $f(U) \subseteq Cl(g^{-1}(W))$. Since g is continuous we have $(gof)(U) \subseteq g(Cl(g^{-1}(W))) \subseteq g(g^{-1}(Cl(W))) \subseteq Cl(W)$. □

THEOREM 2.14. *A function $f : X \to Y$ is weakly ωb-continuous if and only if the graph function $g : X \to X \times Y$ of f defined by $g(x) = (x, f(x))$ for each $x \in X$, is weakly ωb-continuous.*

Proof.

$\Rightarrow$) Suppose that f is weakly ωb-continuous. Let $x \in X$ and W be an open set in $X \times Y$ containing $g(x)$. Then there exists a basic open set $U_1 \times V$ in $X \times Y$ such that $g(x) = (x, f(x)) \in U_1 \times V \subseteq W$. Since f is weakly ωb-continuous there exists an $\omega b - open$ set U_2 in X containing x such that $f(U_2) \subseteq Cl(V)$. Let $U = U_1 \cap U_2$ then U is an $\omega b - open$ set in X with $x \in U$ and $g(U) \subseteq Cl(W)$.

$\Leftarrow$) Suppose that g is weakly ωb-continuous. Let $x \in X$ and V be an open set in Y containing $f(x)$. Then $X \times V$ is an open set containing $g(x)$ and hence there exists an $\omega b - open$ set U in X containing x such that $g(U) \subseteq Cl(X \times V) = X \times Cl(V)$. Therefore, we have $f(U) \subseteq Cl(V)$ and hence f is weakly ωb-continuous. □

THEOREM 2.15. *If $f : (X, \tau) \to (Y, \rho)$ is a weakly ωb-continuous function and Y is Hausdorff then the set $G(f) = \{(x, f(x) : x \in X\}$ is an $\omega b - closed$ set in $X \times Y$.*

Proof. Let $(x, y) \in (X \times Y) - G(f)$. Then $y \neq f(x)$. Since Y is Hausdorff, there exist two disjoint open sets U and V such that $y \in U$ and $f(x) \in V$. Since f is weakly ωb-continuous, there exists an $\omega b - open$ set W containing x such that $f(W) \subseteq Cl(V)$. Since V and U are disjoint, we have $U \cap Cl(V) = \phi$ and hence $U \cap f(W) = \phi$. This shows that $(W \times U) \cap G(f) = \phi$. Then $G(f)$ is $\omega b - closed$. □

THEOREM 2.16. *If $f : X_1 \to Y$ is ωb-continuous, $g : X_2 \to Y$ is weakly ωb-continuous and Y is Hausdorff, then the set $A = \{(x_1, x_2) \in X_1 \times X_2 : f(x_1) = g(x_2)\}$ is $\omega b - closed$ in $X_1 \times X_2$.*

Proof. Let $(x_1, x_2) \in (X_1 \times X_2) - A$. Then $f(x_1) \neq g(x_2)$ and there exist open sets V_1 and V_2 in Y such that $f(x_1) \in V_1$, $g(x_2) \in V_2$ and $V_1 \cap V_2 = \phi$, hence $V_1 \cap Cl(V_2) = \phi$. Since f is ωb-continuous there exists an $\omega b - open$ set U_1 in X_1 containing x_1 such that $f(U_1) \subseteq V_1$. Since g is weakly ωb-continuous there exists an $\omega b - open$ set U_2 in X_2 containing x_2 such that $g(U_2) \subseteq Cl(V_2)$. Now $U_1 \times U_2$ is an $\omega b - open$ set in $X_1 \times X_2$ with $(x_1, x_2) \in U_1 \times U_2 \subseteq (X_1 \times X_2) - A$. This shows that A is $\omega b - closed$ in $X_1 \times X_2$. □

THEOREM 2.17. *If (Y, ρ) is a regular space then a function $f : (X, \tau) \to (Y, \rho)$ is weakly ωb-continuous if and only if it is ωb-continuous*

Proof.

$\Rightarrow$) Let x be any point in X and V be any open set in Y containing $f(x)$. Since (Y, ρ) is regular, there exists $W \in \rho$ such that $f(x) \in W \subseteq Cl(W) \subseteq V$. Since f is weakly ωb-continuous there exists an $\omega b - open$ set U in X containing x such that $f(U) \subseteq Cl(W)$. So $f(U) \subseteq V$. Therefore, f is ωb-continuous.

$\Leftarrow$) Clear. □

DEFINITION 2.18. *Any weakly ωb-continuous function $f : X \to A$, where $A \subseteq X$ and $f_A = f \mid_A$ is the identity function on A, is called weakly ωb-continuous retraction.*

THEOREM 2.19. *Let $f : X \to A$ be a weakly ωb-continuous retraction of X onto A where $A \subseteq X$. If X is a Hausdorff space, then A is an $\omega b - closed$ set in X.*

Proof. Suppoe that A is not $\omega b - closed$ in X. Then there exist a point $x \in \omega bCl(A) - A$. Since f is weakly ωb-continuous retraction, we have $f(x) \neq x$. Since X is Hausdorff, there exist two disjoint open sets U and V such that $x \in U$ and $f(x) \in V$. Then we have $U \cap Cl(V) = \phi$. Now, Let W be any $\omega b - open$ set in X containing x. Then $U \cap W$ is an $\omega b - open$ set containing x and hence $(U \cap W) \cap A \neq \phi$ because $x \in \omega bCl(A)$. Let $y \in (U \cap W) \cap A$. Since $y \in A$, $f(y) = y \in U$ and hence $f(y) \notin Cl(V)$. This gives that $f(W)$ is not a subset of $Cl(V)$. This contradicts the fact that f is weakly ωb-continuous. Therefore A is $\omega b - closed$ in X. □

DEFINITION 2.20. *A space X is called:*

(a) $\omega b - T_1$ if for each pair of distinct points x and y in X, there exist two $\omega b - open$ sets U and V of X containing x and y, respectively, such that $y \notin U$ and $x \notin V$.

(b) $\omega b - T_2$ if for each pair of distinct points x and y in X, there exist two $\omega b - open$ sets U and V of X containing x and y, respectively, such that $U \cap V = \phi$

THEOREM 2.21. *If for each pair of distinct points x and y in a space X there exists a function f of X into a Hausdorff space Y such that*

1) $f(x) \neq f(y)$
2) *f is ωb-continuous at x and*
3) *f is weakly ωb-continuous at y,*

then X is $\omega b - T_2$.

Proof. Since $f(x) \neq f(y)$ and Y is Hausdorff, there exist open sets V_1 and V_2 of Y containing $f(x)$ and $f(y)$, respectively, such that $V_1 \cap V_2 = \phi$, hence $V_1 \cap Cl(V_2) = \phi$. Since f is ωb-continuous at x, there exists an $\omega b-open$ set U_1 in X containing x such that $f(U_1) \subseteq V_1$. Since f is weakly ωb-continuous at y, there exists an $\omega b - open$ set U_2 in X containing y such that $f(U_2) \subseteq Cl(V_2)$. Therefore we obtain $U_1 \cap U_2 = \phi$. This shows that X is $\omega b - T_2$. □

DEFINITION 2.22. *A space X is called Urysohn [5] if for each pair of distinct points x and y in X, there exist open sets U and V such that $x \in U$, $y \in V$ and $Cl(U) \cap Cl(V) = \phi$.*

THEOREM 2.23. *Let $f : (X, \tau) \to (Y, \rho)$ be a weakly ωb-continuous injection. Then the following hold:*

(a) If Y is Hausdorff, then X is $\omega b - T_1$.

(b) If Y is Urysohn, then X is $\omega b - T_2$.

Proof.

(a) Let x_1, $x_2 \in X$ with $x_1 \neq x_2$. Then $f(x_1) \neq f(x_2)$ and there exist open sets V_1 and V_2 in Y containing $f(x_1)$ and $f(x_2)$, respectively, such that $V_1 \cap V_2 = \phi$. Then we obtain $f(x_1) \notin Cl(V_2)$ and $f(x_2) \notin Cl(V_1)$. Since f is weakly ωb-continuous, there exist $\omega b - open$ sets U_1 and U_2 with $x_1 \in U_1$ and $x_2 \in U_2$ such that $f(U_1) \subseteq Cl(V_1)$ and $f(U_2) \subseteq Cl(V_2)$. Hence we obtain $x_2 \notin U_1$ and $x_1 \notin U_2$. This shows that X is $\omega b - T_1$.

(b) Let x_1, $x_2 \in X$ with $x_1 \neq x_2$. Then $f(x_1) \neq f(x_2)$ and there exist open sets V_1 and V_2 in Y containing $f(x_1)$ and $f(x_2)$, respectively, such that $Cl(V_1) \cap Cl(V_2) = \phi$. Since f is weakly ωb-continuous there exist $\omega b-open$ sets U_1 and U_2 in X with $x_1 \in U_1$ and $x_2 \in U_2$ such that $f(U_1) \subseteq Cl(V_1)$ and $f(U_2) \subseteq Cl(V_2)$.Since $f^{-1}(Cl(V_1)) \cap f^{-1}(Cl(V_2)) = \phi$ we obtain $U_1 \cap U_2 = \phi$. Hence X is $\omega b - T_2$. □

DEFINITION 2.24. *A function $f : X \to Y$ is said to have a strongly $\omega b - closed$ graph if for each $(x, y) \in (X \times Y) - G(f)$ there exist an $\omega b-open$ subset U of X and an open subset V of Y such that $(x, y) \in U \times V$ and $(U \times Cl(V)) \cap G(f) = \phi$.*

THEOREM 2.25. *If Y is a Urysohn space and $f : X \to Y$ is weakly ωb-continuous, then $G(f)$ is strongly $\omega b-closed$.*

Proof. Let $(x,y) \in (X \times Y) - G(f)$. Then $y \neq f(x)$ and there exist open sets V and W in Y with $f(x) \in V$ and $y \in W$ such that $Cl(V) \cap Cl(W) = \phi$. Since f is weakly ωb-continuous, there exists an $\omega b - open$ subset U of X containing x such that $f(U) \subseteq Cl(V)$. Therefore we obtain $f(U) \cap Cl(W) = \phi$ and hence $(U \times Cl(W)) \cap G(f) = \phi$. This shows that $G(f)$ is strongly $\omega b - closed$ in $X \times Y$. □

THEOREM 2.26. *Let $f : (X,\tau) \to (Y,\rho)$ be a weakly ωb-continuous function having strongly $\omega b - closed$ graph $G(f)$. If f is injective, then X is $\omega b - T_2$.*

Proof. Let $x_1,\ x_2 \in X$ with $x_1 \neq x_2$. Since f is injective, $f(x_1) \neq f(x_2)$ and $(x_1, f(x_2)) \notin G(f)$. Since $G(f)$ is strongly $\omega b - closed$, there exist an $\omega b - open$ subset U of X containing x_1 and an open subset V of Y such that $(x_1, f(x_2)) \in U \times V$ and $(U \times Cl(V)) \cap G(f) = \phi$ and hence $f(U) \cap Cl(V) = \phi$. Since f is weakly ωb-continuous, there exists an $\omega b - open$ subset W of X containing x_2 such that $f(W) \subseteq Cl(V)$. Therefore, we have $f(U) \cap f(W) = \phi$ and hence $U \cap W = \phi$. This shows that X is $\omega b - T_2$. □

DEFINITION 2.27. *A space X is said to be ωb-connected if X can not be written as a union of two non-empty disjoint $\omega b - open$ sets.*

THEOREM 2.28. *If X is an ωb-connected space and $f : X \to Y$ is weakly ωb-continuous surjection then Y is connected.*

Proof. Suppose that Y is not connected. Then there exist two non-empty disjoint open sets U and V in Y such that $U \cup V = Y$. Hence, we have $f^{-1}(U) \cap f^{-1}(V) = \phi$, $f^{-1}(U) \cup f^{-1}(V) = X$ and since f is surjection we have $f^{-1}(U) \neq \phi \neq f^{-1}(V)$. By Theorem 2.8, we have $f^{-1}(U) \subseteq \omega bInt[f^{-1}(Cl(U))]$ and $f^{-1}(V) \subseteq \omega bInt[f^{-1}(Cl(V))]$. Since U and V are clopen we have $f^{-1}(U) \subseteq \omega bInt[f^{-1}(U)]$ and $f^{-1}(V) \subseteq \omega bInt[f^{-1}(V)]$ and hence $f^{-1}(U)$ and $f^{-1}(V)$ are $\omega b - open$. This implies that X is not ωb-connected which is a contradiction. Therefore Y is connected. □

DEFINITION 2.29. *A topological space (X,τ) is said to be:*

(a) almost compact [3] if every open cover of X has a finite subfamily whose closures cover X.

(b) almost Lindelöf [6] if every open cover of X has a countable subfamily whose closures cover X.

DEFINITION 2.30. *A topological space (X,τ) is said to be ωb-compact (resp. ωb-Lindelöf) if every $\omega b - open$ cover of X has a finite (resp. countable) subcover.*

THEOREM 2.31. *Let $f : X \to Y$ be a weakly ωb-continuous surjection. Then the following hold:*

(a) If X is ωb-compact, then Y is almost compact.

(b) If X is ωb-Lindelöf, then Y is almost Lindelöf.

Proof.

(a) Let $\{V_\alpha : \alpha \in \Delta\}$ be a cover of Y by open sets in Y. For each $x \in X$ there exists $V_{\alpha_x} \in \{V_\alpha : \alpha \in \Delta\}$ such that $f(x) \in V_{\alpha_x}$. Since f is weakly ωb-continuous, there exists an $\omega b - open$ set U_x of X containing x such that $f(U_x) \subseteq Cl(V_{\alpha_x})$. The family $\{U_x : x \in X\}$ is a cover of X by $\omega b - open$ sets of X and hence there exists a finite subset X_0 of X such that $X \subseteq \cup\{U_x : x \in X_0\}$. Therefore, we obtain $Y = f(X) \subseteq \cup\{Cl(V_{\alpha_x}) : x \in X_0\}$. This shows that Y is almost compact.

(b) Similar to *(a)*. □

REFERENCES

[1] D. ANDRIJEVIĆ, *On b-open sets*, Mat. Vesnik **48** (1996), 59–64.

[2] A.A. EL-ATIK, *A study on some types of mappings on topological spaces*, M.Sc. Thesis, Tanta University, Egypt (1997).

[3] T. NOIRI, *Between continuity and weak continuity*, Boll. Unione Mat. Ital. **9** (1974), 647–654.

[4] T. NOIRI, A. AL-OMARI AND M.S.M. NOORANI, *On ωb-open sets and b-Lindelöf spaces*, Eur. J. Pure Appl. Math. **1** (2008), 3–9.

[5] S. WILLARD, *General topology*, Addison Wesley, London (1970).

[6] S. WILLARD AND U.N.B. DISSANAYAKE, *The almost Lindelöf degree*, Canadian Math. Bull. **27** (1984), 452–455.

[7] U. SENGUL, *Weakly b-continuous functions*, Chaos Solitons Fractals **41** (2009), 1070–1077.

Author's address:

Jamal M. Mustafa
Faculty of Science, Department of Sciences
Jerash University
Jerash, Jordan
E-mail: jjmmrr971@yahoo.com

Received February 6, 2011
Revised March 21, 2011

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 61–94

Vector Bundles on Elliptic Curves and Factors of Automorphy

Oleksandr Iena

Communicated by Ugo Bruzzo

Abstract. *We translate Atiyah's results on classification of vector bundles on elliptic curves to the language of factors of automorphy.*

Keywords: Elliptic Curves, Factors of Automorphy, Theta Functions, Vector Bundles.
MS Classification 2010: 14H60, 14H52, 14D20, 14K25

1. Introduction

1.1. Motivation

The problem of classification of vector bundles over an elliptic curve was considered and completely solved by Atiyah in [1].

For a group Γ acting on a complex manifold Y, an r-dimensional factor of automorphy is a holomorphic function $f : \Gamma \times Y \to \mathrm{GL}_r(\mathbb{C})$ satisfying $f(\lambda\mu, y) = f(\lambda, \mu y) f(\mu, y)$. Two factors of automorphy f and f' are equivalent if there exists a holomorphic function $h : Y \to \mathrm{GL}_r(\mathbb{C})$ such that $h(\lambda y) f(\lambda, y) = f'(\lambda, y) h(y)$.

Given a complex manifold X and the universal covering $Y \xrightarrow{p} X$, let Γ be the fundamental group of X acting naturally on Y by deck transformations. Then there is a one-to-one correspondence between equivalence classes of r-dimensional factors of automorphy and isomorphism classes of vector bundles on X with trivial pull-back along p. In particular, if Y does not possess any non-trivial vector bundles, one obtains a one-to-one correspondence between equivalence classes of r-dimensional factors of automorphy and isomorphism classes of vector bundles on X. In particular this is the case for complex tori.

Since it is known that one-dimensional complex tori correspond to elliptic curves and since the classification of holomorphic vector bundles on a projective variety over $\mathbb{C}$ is equivalent to the classification of algebraic vector bundles (cf. [13]), it is possible to formulate Atiyah's results in the language of factors of automorphy. So for example in the case of vector bundles of rank 1 and 2 such a formulation using factors of automorphy was given in [4], Theorems 4.4 and 4.5.

This paper is a shortened version of the diploma thesis [7] and aims to prove some results used without any proofs by different authors, in particular in [12] and [3]. The main result of this note, Theorem 5.24, gives a classification of indecomposable vector bundles of fixed rank and degree on a complex torus in terms of factors of automorhy. Its statement coincides with the statement of Proposition 1 from [12], which was given without any proof.

The author thanks Igor Burban, Bernd Kreußler, and Günther Trautmann, who motivated him to prepare this manuscript.

1.2. Structure of the Paper

In Section 2 we establish a correspondence between vector bundles and factors of automorphy. Section 3 deals with properties of factors of automorphy, in particular we discuss a correspondence between operations on vector bundles and operations on factors of automorphy. From Section 4 on we restrict ourselves to the case of vector bundles on complex tori. It is shown in Theorem 4.11 that to define a vector bundle of rank r on a complex one-dimensional torus is the same as to fix a holomorphic function $\mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$. In Section 5 we first present in Theorem 5.13 a classification of indecomposable vector bundles of degree zero, using this we give then in Theorem 5.24 a complete classification of indecomposable vector bundles of fixed rank and degree in terms of factors of automorphy.

1.3. Notations and Conventions

Following Atiyah's paper [1] we denote by $\mathcal{E}(r,d) = \mathcal{E}_X(r,d)$ the set of isomorphism classes of indecomposable vector bundles over X of rank r and degree d. For a vector bundle E we usually denote the corresponding locally free sheaf of its sections by $\mathcal{E}$. By Vect we denote the category of finite dimensional vector spaces. For a divisor D we denote by $[D]$ the corresponding line bundle.

2. Correspondence between Vector Bundles and Factors of Automorphy

Let X be a complex manifold and let $p : Y \to X$ be a covering of X. Let $\Gamma \subset \mathrm{Deck}(Y/X)$ be a subgroup in the group of deck transformations $\mathrm{Deck}(Y/X)$ such that for any two points y_1 and y_2 with $p(y_1) = p(y_2)$ there exists an element $\gamma \in \Gamma$ such that $\gamma(y_1) = y_2$. In other words, Γ acts transitively in each fiber. We call this property $(\mathbf{T})$.

Remark 2.1. *Note that for any two points* y_1 *and* y_2 *there can be only one* $\gamma \in \mathrm{Deck}(Y/X)$ *with* $\gamma(y_1) = y_2$ *(see [5], Satz 4.8). Therefore,* $\Gamma = \mathrm{Deck}(Y/X)$ *and the property (**T**) simply means that* $p : Y \to X$ *is a normal (Galois) covering.*

We have an action of Γ on Y:

$$\Gamma \times Y \to Y, \quad y \mapsto \gamma(y) =: \gamma y.$$

Definition 2.2. *A holomorphic function* $f : \Gamma \times Y \to \mathrm{GL}_r(\mathbb{C})$, $r \in \mathbb{N}$ *is called an* r*-dimensional factor of automorphy if it satisfies the relation*

$$f(\lambda\mu, y) = f(\lambda, \mu y) f(\mu, y).$$

Denote by $Z^1(\Gamma, r)$ *the set of all* r*-dimensional factors of automorphy.*

We introduce the relation $\sim$ on $Z^1(\Gamma, r)$. We say that f is equivalent to f' if there exists a holomorphic function $h : Y \to \mathrm{GL}_r(\mathbb{C})$ such that

$$h(\lambda y) f(\lambda, y) = f'(\lambda, y) h(y).$$

We write in this case $f \sim f'$.

Lemma 2.3. *The relation* $\sim$ *is an equivalence relation on* $Z^1(\Gamma, r)$.

Proof. Straightforward verifications. $\square$

We denote the set of equivalence classes of $Z^1(\Gamma, r)$ with respect to $\sim$ by $H^1(\Gamma, r)$.

Consider $f \in Z^1(\Gamma, r)$ and a trivial vector bundle $Y \times \mathbb{C}^r \to Y$. Define a holomorphic action of Γ on $Y \times \mathbb{C}^r$:

$$\Gamma \times Y \times \mathbb{C}^r \to Y \times \mathbb{C}^r, \quad (\lambda, y, v) \mapsto (\lambda y, f(\lambda, y)v) =: \lambda(y, v).$$

Denote $E(f) = Y \times \mathbb{C}^r/\Gamma$ and note that for two equivalent points $(y, v) \sim_\Gamma (y', v')$ with respect to the action of Γ on $Y \times \mathbb{C}^r$ it follows that $p(y) = p(y')$. In fact, $(y, v) \sim_\Gamma (y', v')$ implies in particular that $y = \gamma y'$ for some $\gamma \in \Gamma$ and by the definition of deck transformations $p(y) = p(\gamma y') = p(y')$. Hence the projection $Y \times \mathbb{C}^r \to Y$ induces the map

$$\pi : E(f) \to X, \quad [y, v] \mapsto p(y).$$

We equip $E(f)$ with the quotient topology.

Theorem 2.4. $E(f)$ *inherits a complex structure from* $Y \times \mathbb{C}^r$ *and the map* $\pi : E(f) \to X$ *is a holomorphic vector bundle on* X.

Proof. First we prove that π is a topological vector bundle. Clearly π is a continuous map. Consider the commutative diagram

$$\begin{tikzcd} Y \times \mathbb{C}^r \arrow[r] \arrow[d] & E(f) \arrow[d, "\pi"] \\ Y \arrow[r, "p"] & X. \end{tikzcd}$$

Let x be a point of X. Since p is a covering, one can choose an open neighbourhood U of x such that its preimage is a disjoint union of open sets biholomorphic to U, i. e., $p^{-1}(U) = \bigsqcup_{i\in\mathcal{I}} V_i$, $p_i := p|_{V_i} : V_i \to U$ is a biholomorphism for each $i \in \mathcal{I}$. For each pair $(i,j) \in \mathcal{I}\times\mathcal{I}$ there exists a unique $\lambda_{ij} \in \Gamma$ such that $\lambda_{ij} p_j^{-1}(x) = p_i^{-1}(x)$ for all $x \in U$. This follows from the property (**T**).

We have $\pi^{-1}(U) = ((\bigsqcup_{i\in\mathcal{I}} V_i) \times \mathbb{C}^r)/\Gamma$.

Choose some $i_U \in \mathcal{I}$. Consider the holomorphic map

$$\varphi'_U : \left(\bigsqcup_{i\in\mathcal{I}} V_i\right) \times \mathbb{C}^r \to U \times \mathbb{C}^r, \quad (y_i, v) \mapsto (p(y_i), f(\lambda_{i_U i}, y_i)v),\ y_i \in V_i.$$

Suppose that $(y_i, v') \sim_\Gamma (y_j, v)$. This means

$$(y_i, v') = \lambda_{ij}(y_j, v) = (\lambda_{ij} y_j, f(\lambda_{ij}, y_j)v).$$

Therefore,

$$\begin{aligned}\varphi'_U(y_i, v') &= (p(y_i), f(\lambda_{i_U i}, y_i)v') = (p(\lambda_{ij} y_j), f(\lambda_{i_U i}, \lambda_{ij} y_j) f(\lambda_{ij}, y_j)v) \\ &= (p(y_j), f(\lambda_{i_U j}, y_j)v) = \varphi'_U(y_j, v).\end{aligned}$$

Thus φ'_U factorizes through $((\bigsqcup_{i\in\mathcal{I}} V_i) \times \mathbb{C}^r)/\Gamma$, i. e., the map

$$\varphi_U : \left(\left(\bigsqcup_{i\in\mathcal{I}} V_i\right) \times \mathbb{C}^r\right)/\Gamma \to U \times \mathbb{C}^r, \quad [(y_i, v)] \mapsto (p(y_i), f(\lambda_{i_U i}, y_i)v),\ y_i \in V_i$$

is well-defined and continuous. We claim that φ_U is bijective.

Suppose $\varphi_U([(y_i, v')]) = \varphi_U([(y_j, v)])$, where $y_i \in V_i$, $y_j \in V_j$. By definition this is equivalent to $(p(y_i), f(\lambda_{i_U i}, y_i)v') = (p(y_j), f(\lambda_{i_U j}, y_j)v)$, which means $y_i = \lambda_{ij} y_j$ and

$$\begin{aligned}f(\lambda_{i_U i}, \lambda_{ij} y_j)v' &= f(\lambda_{i_U i}, y_i)v' = f(\lambda_{i_U j}, y_j)v \\ &= f(\lambda_{i_U i}\lambda_{ij}, y_j)v = f(\lambda_{i_U i}, \lambda_{ij} y_j) f(\lambda_{ij}, y_j)v.\end{aligned}$$

We conclude $v' = f(\lambda_{ij}, y_j)v$ and $[(y_i, v')] = [(y_j, v)]$, which means injectivity of φ_U.

At the same time for each element $(y, v) \in U \times \mathbb{C}^r$ one has

$$\begin{aligned}&\varphi_U([(p_i^{-1}(y), f(\lambda_{i_U i}, p_i^{-1}(y))^{-1}v)])\\ &\qquad = (pp_i^{-1}(y), f(\lambda_{i_U i}, p_i^{-1}(y))f(\lambda_{i_U i}, p_i^{-1}(y))^{-1}v) = (y, v),\end{aligned}$$

i.e., φ_U is surjective and we obtain that φ_U is a bijective map.

This means, that φ_U is a trivialization for U and that $\pi : E(f) \to X$ is a (continuous) vector bundle. If U and V are two neighbourhoods of X defined as above for which $E(f)|_U$, $E(f)|_V$ are trivial, then the corresponding transition function is

$$\varphi_U\varphi_V^{-1} : (U \cap V) \times \mathbb{C}^r \to (U \cap V) \times \mathbb{C}^r, \quad (x, v) \mapsto (x, g_{UV}(x)v),$$

where $g_{UV} : U \cap V \to \mathrm{GL}_r(\mathbb{C})$ is a cocycle defining $E(f)$. But from the construction of φ_U it follows that

$$g_{UV}(x) = f(\lambda_{i_U i_V}, p_{i_V}^{-1}(x)).$$

Therefore, g_{UV} is a holomorphic map, hence $\varphi_U\varphi_V^{-1}$ is also a holomorphic map. Thus the maps φ_U give $E(f)$ a complex structure. Since π is locally a projection, one sees that π is a holomorphic map. □

Remark 2.5. *Note that $p^*E(f)$ is isomorphic to $Y \times \mathbb{C}^r$. An isomorphism can be given by the map*

$$p^*E(f) \to Y \times \mathbb{C}^r, \quad (y, [\tilde{y}, v]) \mapsto (y, f(\lambda, \tilde{y})v), \quad \lambda\tilde{y} = y.$$

Now we have the map from $Z^1(\Gamma, r)$ to the set $K_r = \{[E] \mid p^*(E) \simeq Y \times \mathbb{C}^r\}$ of isomorphism classes of vector bundles of rank r over X with trivial pull back with respect to p.

$$\phi' : Z^1(\Gamma, r) \to K_r; \quad f \mapsto [E(f)].$$

Theorem 2.6. *Let K_r denote the set of isomorphism classes of vector bundles of rank r on X with trivial pull back with respect to p. Then the map*

$$H^1(\Gamma, r) \to K_r, \quad [f] \mapsto [E(f)]$$

is a bijection.

Proof. This proof generalizes the proof from [2, Appendix B] given only for line bundles.

Consider the map $\phi' : Z^1(\Gamma, r) \to K_r$ and let f and f' be two equivalent r-dimensional factors of automorphy. It means that there exists a holomorphic function $h : Y \to \mathrm{GL}_r(\mathbb{C})$ such that

$$f'(\lambda, y) = h(\lambda y)f(\lambda, y)h(y)^{-1}.$$

Therefore, for two neighbourhoods U, V constructed as above we have the following relation for cocycles corresponding to f and f'.

$$\begin{aligned} g'_{UV}(x) &= f'(\lambda_{UV}, p_{i_V}^{-1}(x)) = h(\lambda_{UV} p_{i_V}^{-1}(x)) f(\lambda_{UV}, p_{i_V}^{-1}(x)) h(p_{i_V}^{-1}(x))^{-1} \\ &= h(p_{i_U}^{-1}(x)) g_{UV}(x) h(p_{i_V}^{-1}(x))^{-1} = h_U(x) g_{UV}(x) h_V(x)^{-1}, \end{aligned}$$

where $\lambda_{UV} = \lambda_{i_U i_V}$, $h_U(x) = h(p_{i_U}^{-1}(x))$ and $h_V(x) = h(p_{i_V}^{-1}(x))$. We obtained

$$g'_{UV} = h_U g_{UV} h_U^{-1},$$

which is exactly the condition for two cocycles to define isomorphic vector bundles. Therefore, $E(f) \simeq E(f')$ and it means that ϕ' factorizes through $H^1(\Gamma, r)$, i.e., the map

$$\phi : H^1(\Gamma, r) \to K_r; \quad [f] \mapsto [E(f)]$$

is well-defined.

It remains to construct the inverse map. Suppose $E \in K_r$, in other words $p^*(E)$ is the trivial bundle of rank r over Y. Let $\alpha : p^*E \to Y \times \mathbb{C}^r$ be a trivialization. The action of Γ on Y induces a holomorphic action of Γ on p^*E :

$$\lambda(y, e) := (\lambda y, e) \text{ for } (y, e) \in p^*E = Y \times_X E.$$

Via α we get for every $\lambda \in \Gamma$ an automorphism ψ_λ of the trivial bundle $Y \times \mathbb{C}^r$. Clearly ψ_λ should be of the form

$$\psi_\lambda(y, v) = (\lambda y, f(\lambda, y)v),$$

where $f : \Gamma \times Y \to \mathrm{GL}_r(\mathbb{C})$ is a holomorphic map. The equation for the action $\psi_{\lambda\mu} = \psi_\lambda \psi_\mu$ implies that f should be an r-dimensional factor of automorphy.

Suppose α' is an another trivialization of p^*E. Then there exists a holomorphic map $h : Y \to \mathrm{GL}_r(\mathbb{C})$ such that $\alpha'\alpha^{-1}(y, v) = (y, h(y)v)$. Let f' be a factor of automorphy corresponding to α'. From

$$\begin{aligned} (\lambda y, f'(\lambda, y)v) &= \psi'_\lambda(y, v) = \alpha'\lambda\alpha'^{-1}(y, v) = \alpha'\alpha^{-1}\alpha\lambda\alpha^{-1}\alpha\alpha'^{-1}(y, v) \\ &= \alpha'\alpha^{-1}\psi_\lambda(\alpha'\alpha^{-1})^{-1}(y, v) = \alpha'\alpha^{-1}\psi_\lambda(y, h(y)^{-1}v) \\ &= \alpha'\alpha^{-1}(\lambda y, f(\lambda, y)h(y)^{-1}v) = (\lambda y, h(\lambda y)f(\lambda, y)h(y)^{-1}), \end{aligned}$$

we obtain $f'(\lambda, y) = h(\lambda y)f(\lambda, y)h(y)^{-1}$. The last means that $[f] = [f']$, in other words, the class of a factor of automorphy in $H^1(\Gamma, r)$ does not depend on the trivialization and we get a map $K_r \to H^1(\Gamma, r)$. This map is the inverse of ϕ. □

Let X be a connected complex manifold, let $p : \tilde{X} \to X$ be a universal covering of X, and let $\Gamma = \mathrm{Deck}(Y/X)$. Since universal coverings are normal coverings, Γ satisfies the property (**T**) (see [5, Satz 5.6]). Moreover, Γ is isomorphic to the fundamental group $\pi_1(X)$ of X (see [5, Satz 5.6]). An isomorphism is given as follows.

Fix $x_0 \in X$ and $\tilde{x}_0 \in \tilde{X}$ with $p(\tilde{x}_0) = x_0$. We define a map

$$\Phi : \mathrm{Deck}(\tilde{X}/X) \to \pi_1(X, x_0)$$

as follows. Let $\sigma \in \mathrm{Deck}(\tilde{X}/X)$ and $v : [0;1] \to \tilde{X}$ be a curve with $v(0) = \tilde{x}_0$ and $v(1) = \sigma(\tilde{x}_0)$. Then a curve

$$pv : [0;1] \to X, \quad t \mapsto pv(t)$$

is such that $pv(0) = pv(1) = x_0$. Define $\Phi(\sigma) := [pv]$, where $[pv]$ denotes a homotopy class of pv. The map Φ is well defined and is an isomorphism of groups.

So we can identify Γ with $\pi_1(X)$. Therefore, we have an action of $\pi_1(X)$ on $\tilde{X}$ by deck transformations.

Consider an element $[w] \in \pi_1(X, x_0)$ represented by a path $w : [0;1] \to X$. We denote $\sigma = \Phi^{-1}([w])$. Consider any $\tilde{x}_0 \in X$ such that $p(\tilde{x}_0) = w(0) = w(1)$, then the path w can be uniquely lifted to the path

$$v : [0;1] \to \tilde{X}$$

with $v(0) = \tilde{x}_0$ (see [5], Satz 4.14). Denote $\tilde{x}_1 = v(1)$. Then σ is a unique element in $\mathrm{Deck}(\tilde{X}/X)$ such that $\sigma(\tilde{x}_0) = \tilde{x}_1$. This gives a description of the action of $\pi_1(X, x_0)$ on $\tilde{X}$.

Now we have a corollary to Theorem 2.6.

Corollary 2.7. *Let X be a connected complex manifold, let $p : \tilde{X} \to X$ be the universal covering, let Γ be the fundamental group of X naturally acting on $\tilde{X}$ by deck transformations. As above, $H^1(\Gamma, r)$ denotes the set of equivalence classes of r-dimensional factors of automorphy*

$$\Gamma \times \tilde{X} \to \mathrm{GL}_r(\mathbb{C}).$$

Then there is a bijection

$$H^1(\Gamma, r) \to K_r, \quad [f] \mapsto E(f),$$

where K_r denotes the set of isomorphism classes of vector bundles of rank r on X with trivial pull back with respect to p.

3. Properties of Factors of Automorphy

DEFINITION 3.1. *Let $f : \Gamma \times Y \to \mathrm{GL}_r(\mathbb{C})$ be an r-dimensional factor of automorphy. A holomorphic function $s : Y \to \mathbb{C}^r$ is called an f-theta function if it satisfies*

$$s(\gamma y) = f(\gamma, y)s(y) \text{ for all } \gamma \in \Gamma,\ y \in Y.$$

THEOREM 3.2. *Let $f : \Gamma \times Y \to \mathrm{GL}_r(\mathbb{C})$ be an r-dimensional factor of automorphy. Then there is a one-to-one correspondence between sections of $E(f)$ and f-theta functions.*

Proof. Let $\{V_i\}_{i\in\mathcal{I}}$ be a covering of Y such that p restricted to V_i is a homeomorphism. Denote $\varphi_i := (p|_{V_i})^{-1}$, $U_i := p(V_i)$. Then $\{U_i\}$ is a covering of X such that $E(f)$ is trivial over each U_i.

Consider a section of $E(f)$ given by functions $s_i : U_i \to \mathbb{C}^r$ satisfying

$$s_i(x) = g_{ij}(x)s_j(x) \text{ for } x \in U_i \cap U_j,$$

where

$$g_{ij}(x) = f(\lambda_{U_iU_j}, \varphi_j(x)), \quad x \in U_i \cap U_j$$

is a cocycle defining $E(f)$ (see the proof of Theorem 2.6). Define $s : Y \to \mathbb{C}^r$ by $s(\varphi_i(x)) := s_i(x)$. To prove that this is well-defined we need to show that $s_i(x) = s_j(x)$ when $\varphi_i(x) = \varphi_j(x)$. But since $\varphi_i(x) = \varphi_j(x)$ we obtain $\lambda_{U_iU_j} = 1$. Therefore,

$$s_i(x) = g_{ij}(x)s_i(x) = f(\lambda_{U_iU_j}, \varphi_j(x))s_j(x) = f(1, \varphi_j(x))s_j(x) = s_j(x).$$

For any $\gamma \in \Gamma$ for any point $y \in Y$ take $i, j \in \mathcal{I}$ and $x \in X$ such that $y = \varphi_j(x)$ and $\gamma y = \gamma\varphi_j(x) = \varphi_i(x)$. Thus $\gamma = \lambda_{U_iU_j}$ and one obtains

$$\begin{aligned} s(\gamma y) &= s(\varphi_i(x)) = s_i(x) = g_{ij}(x)s_j(x) \\ &= f(\lambda_{U_iU_j}, \varphi_j(x))s_j(x) = f(\gamma, y)s(\varphi_j(x)) = f(\gamma, z)s(y). \end{aligned}$$

In other words, s is an f-theta function.

Vice versa, let $s : Y \to \mathbb{C}^r$ be an f-theta function. We define $s_i : U_i \to \mathbb{C}^r$ by $s_i(x) := s(\varphi_i(x))$. Then for a point $x \in U_i \cap U_j$ we have

$$\begin{aligned} s_i(x) &= s(\varphi_i(x)) = s(\lambda_{U_iU_j}\varphi_j(x)) \\ &= f(\lambda_{U_iU_j}, \varphi_j(x))s(\varphi_j(x)) = g_{ij}(x)s_j(x), \end{aligned}$$

which means that the functions s_i define a section of $E(f)$. The described correspondences are clearly inverse to each other. $\square$

The following statement will be useful in the sequel.

THEOREM 3.3. *Let*

$$f(\lambda, y) = \begin{pmatrix} f'(\lambda, y) & \tilde{f}(\lambda, y) \\ 0 & f''(\lambda, y) \end{pmatrix}$$

be an $r' + r''$-dimensional factor of automorphy, where $f'(\lambda, y) \in \mathrm{GL}_{r'}(\mathbb{C})$, $f''(\lambda, y) \in \mathrm{GL}_{r''}(\mathbb{C})$. Then

(a) $f' : \Gamma \times Y \to \mathrm{GL}_{r'}(\mathbb{C})$ and $f'' : \Gamma \times Y \to \mathrm{GL}_{r''}(\mathbb{C})$ are r' and r''-dimensional factors of automorphy respectively;

(b) there is an extension of vector bundles

$$0 \longrightarrow E(f') \xrightarrow{\ i\ } E(f) \xrightarrow{\ \pi\ } E(f'') \longrightarrow 0\ .$$

Proof. The statement of (a) follows from straightforward verification. To prove (b) we define maps i and π as follows.

$$\begin{aligned} i :& E(f') \to E(f), & [y, v] \mapsto [y, \begin{pmatrix} v \\ 0 \end{pmatrix}], \quad & v \in \mathbb{C}^{r'}, \quad \begin{pmatrix} v \\ 0 \end{pmatrix} \in \mathbb{C}^{r'+r''} \\ \pi :& E(f) \to E(f''), & [y, \begin{pmatrix} v \\ w \end{pmatrix}] \to [y, w], \quad & v \in \mathbb{C}^{r'}, \quad w \in \mathbb{C}^{r''} \end{aligned}$$

Since $[\lambda y, f'(\lambda, y)v]$ is mapped via i to

$$\left[\lambda y, \begin{pmatrix} f'(\lambda, y)v \\ 0 \end{pmatrix}\right] = \left[\lambda y, f(\lambda, y) \begin{pmatrix} v \\ 0 \end{pmatrix}\right],$$

one concludes that i is well-defined. Analogously, since $[\lambda y, f''(\lambda, y)w] = [y, w]$ one sees that π is well-defined. Using the charts from the proof of (2.4) one easily sees that the defined maps are holomorphic.

Notice that i and π respect fibers, i is injective and π is surjective in each fiber. This proves the statement. □

Now we recall one standard construction from linear algebra. Let A be an $m \times n$ matrix. It represents some morphism $\mathbb{C}^n \to \mathbb{C}^m$ for fixed standard bases in $\mathbb{C}^n$ and $\mathbb{C}^m$.

Let $\mathcal{F} : \mathsf{Vect}^p \to \mathsf{Vect}$ be a covariant functor. Let $A_1, \dots, A_p$ be the matrices representing morphisms $\mathbb{C}^n_1 \xrightarrow{f_1} \mathbb{C}^m_1, \dots, \mathbb{C}^n_p \xrightarrow{f_p} \mathbb{C}^m_p$ in standard bases.

If for each object $\mathcal{F}(\mathbb{C}^m)$ we fix some basis, then the matrix corresponding to the morphism $\mathcal{F}(f_1, \dots, f_p)$ is denoted by $\mathcal{F}(A_1, \dots, A_p)$. Clearly it satisfies

$$\mathcal{F}(A_1 B_1, \dots, A_p B_p) = \mathcal{F}(A_1, \dots, A_p)\mathcal{F}(B_1, \dots, B_p).$$

In this way $A \otimes B$, $S^q(A)$, $\Lambda^q(A)$ can be defined. As $\mathcal{F}$ one considers the functors

$$\_ \otimes \_ : \mathsf{Vect}^2 \to \mathsf{Vect}, \quad S^n : \mathsf{Vect} \to \mathsf{Vect}, \quad \Lambda : \mathsf{Vect} \to \mathsf{Vect}$$

respectively.

Recall that every holomorphic functor $\mathcal{F} : \mathsf{Vect}^n \to \mathsf{Vect}$ can be canonically extended to the category of vector bundles of finite rank over X. By abuse of notation we will denote the extended functor by $\mathcal{F}$ as well.

Theorem 3.4. *Let $\mathcal{F} : \mathsf{Vect}^n \to \mathsf{Vect}$ be a covariant holomorphic functor. Let $f_1, \dots, f_n$ be r_i-dimensional factors of automorphy. Then $f = \mathcal{F}(f_1, \dots, f_n)$ is a factor of automorphy defining $\mathcal{F}(E(f_1), \dots, E(f_n))$.*

Proof. One clearly has

$$\begin{aligned}\mathcal{F}(f_1, \dots, f_n)(\lambda\mu, y) &= \mathcal{F}(f_1(\lambda\mu, y), \dots, f_n(\lambda\mu, y)) \\ &= \mathcal{F}(f_1(\lambda, \mu y)f_1(\mu, y), \dots, f_n(\lambda, \mu y)f_n(\mu, y)) \\ &= F(f_1(\lambda, \mu y), \dots, f_n(\lambda, \mu y))F(f_1(\mu, y), \dots, f_n(\mu, y)) \\ &= \mathcal{F}(f_1, \dots, f_n)(\lambda, \mu y)\mathcal{F}(f_1, \dots, f_n)(\mu, y).\end{aligned}$$

Since $(f_1, \dots, f_n)$ represents an isomorphism in Vect^n, $\mathcal{F}(f_1, \dots, f_n)$ also represents an isomorphism $\mathbb{C}^r \to \mathbb{C}^r$ for some $r \in \mathbb{N}$. Therefore, f is an r-dimensional factor of automorphy.

Since $f = \mathcal{F}(f_1, \dots, f_n)$, for cocycles defining the corresponding vector bundles the equality $g_{U_1U_2} = \mathcal{F}(g_{1U_1U_2}, \dots, g_{nU_1U_2})$ holds true, where $g_{iU_1U_2}$ is a cocycle defining $E(f_i)$. This shows that $E(f) = \mathcal{F}(E(f_1), \dots, E(f_n))$ and proves the required statement. $\square$

For example for $\mathcal{F} = \_ \otimes \_ : \mathsf{Vect}^2 \to \mathsf{Vect}$ we get the following obvious corollary.

Corollary 3.5. *Let $f' : \Gamma \times Y \to \mathrm{GL}_{r'}(\mathbb{C})$ and $f'' : \Gamma \times Y \to \mathrm{GL}_{r''}(\mathbb{C})$ be two factors of automorphy. Then $f = f' \otimes f'' : \Gamma \times Y \to \mathrm{GL}_{r'r''}(\mathbb{C})$ is also a factor of automorphy. Moreover, $E(f) \simeq E(f') \otimes E(f'')$.*

It is not essential that the functor in Theorem 3.4 is covariant. The following theorem is a generalization of Theorem 3.4.

Theorem 3.6. *Let $\mathcal{F} : \mathsf{Vect}^n \to \mathsf{Vect}$ be a holomorphic functor. Let $\mathcal{F}$ be covariant in k first variables and contravariant in $n-k$ last variables. Let $f_1, \dots, f_n$ be r_i-dimensional factors of automorphy. Then*

$$f = \mathcal{F}(f_1, \dots, f_k, f_{k+1}^{-1}, \dots, f_n^{-1})$$

is a factor of automorphy defining $\mathcal{F}(E(f_1), \dots, E(f_n))$.

Proof. The proof is analogous to the proof of Theorem 3.4. $\square$

4. Vector Bundles on Complex Tori

4.1. One Dimensional Complex Tori

Let X be a complex torus, i.e., $X = \mathbb{C}/\Gamma$, $\Gamma = \mathbb{Z}\tau + \mathbb{Z}$, $\operatorname{Im}\tau > 0$. Then the universal covering is $\tilde{X} = \mathbb{C}$, namely

$$\mathrm{pr} : \mathbb{C} \to \mathbb{C}/\Gamma, \quad x \mapsto [x].$$

We have an action of Γ on $\mathbb{C}$:

$$\Gamma \times \mathbb{C} \to \mathbb{C}, \quad (\gamma, y) \mapsto \gamma + y.$$

Clearly Γ acts on $\mathbb{C}$ by deck transformations and satisfies the property (**T**).

Since $\mathbb{C}$ is a non-compact Riemann surface, by [5, Theorem 30.4, p. 204], there are only trivial bundles on $\mathbb{C}$. Therefore, we have a one-to-one correspondence between classes of isomorphism of vector bundles of rank r on X and equivalence classes of factors of automorphy

$$f : \Gamma \times \mathbb{C} \to \mathrm{GL}_r(\mathbb{C}).$$

As usually, V_a denotes the standard parallelogram constructed at point a, U_a is the image of V_a under the projection, $\varphi_a : U_a \to V_a$ is the local inverse of the projection.

REMARK 4.1. *Let f be an r-dimensional factor of automorphy. Then*

$$g_{ab}(x) = f(\varphi_a(x) - \varphi_b(x), \varphi_b(x))$$

is a cocycle defining $E(f)$. This follows from the construction of the cocycle in the proof of Theorem 2.6.

EXAMPLE 4.2. *There are factors of automorphy corresponding to classical theta functions. For any theta-characteristic $\xi = a\tau + b$, where $a, b \in \mathbb{R}$, there is a holomorphic function $\theta_\xi : \mathbb{C} \to \mathbb{C}$ defined by*

$$\theta_\xi(z) = \sum_{n\in\mathbb{Z}} \exp(\pi i(n + a)^2\tau)\exp(2\pi i(n + a)(z + b)),$$

which satisfies

$$\theta_\xi(\gamma + z) = \exp(2\pi ia\gamma - \pi ip^2\tau - 2\pi ip(z + \xi))\theta_\xi(z) = e_\xi(\gamma, z)\theta_\xi(z),$$

where $\gamma = p\tau + q$ and $e_\xi(\gamma, z) = \exp(2\pi ia\gamma - \pi ip^2\tau - 2\pi ip(z + \xi))$. Since

$$e_\xi(\gamma_1 + \gamma_2, z) = e_\xi(\gamma_1, \gamma_2 + z)e_\xi(\gamma_2, z),$$

we conclude that $e_\xi(\gamma, z)$ is a factor of automorphy.

By Theorem 3.2 $\theta_\xi(z)$ defines a section of $E(e_\xi(\gamma, z))$.

For more information on classical theta functions see [8, 9, 10].

THEOREM 4.3. $\deg E(e_\xi) = 1$.

Proof. We know that sections of $E(e_\xi)$ correspond to e_ξ - theta functions. The classical e_ξ-theta function $\theta_\xi(z)$ defines a section s_ξ of $E(e_\xi)$. Since θ_ξ has only simple zeros and the set of zeros of $\theta_\xi(z)$ is $\frac{1}{2} + \frac{\tau}{2} + \xi + \Gamma$, we conclude that s_ξ has exactly one zero at point $p = [\frac{1}{2} + \frac{\tau}{2} + \xi] \in X$. Hence by [6, p. 136] we get $E(e_\xi) \simeq [p]$ and thus $\deg E(e_\xi) = 1$. □

THEOREM 4.4. *Let ξ and η be two theta-characteristics. Then*

$$E(e_\xi) \simeq t^*_{[\eta-\xi]} E(e_\eta),$$

where $t_{[\eta-\xi]} : X \to X, \quad x \mapsto x + [\eta - \xi]$ *is the translation by* $[\eta - \xi]$.

Proof. As in the proof of Theorem 4.3 $E(e_\xi) \simeq [p]$ and $E(e_\eta) = [q]$ for $p = [\frac{1}{2} + \frac{\tau}{2} + \xi]$ and $q = [\frac{1}{2} + \frac{\tau}{2} + \eta]$. Since $t_{[\eta-\xi]} p = q$, we get

$$E(e_\xi) \simeq [p] \simeq t^*_{[\eta-\xi]}[q] \simeq t^*_{[\eta-\xi]} E(e_\eta),$$

which completes the proof. □

Now we are going to investigate the extensions of the type

$$0 \to X \times \mathbb{C} \to E \to X \times \mathbb{C} \to 0.$$

In this case the transition functions are given by matrices of the type

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

and E is isomorphic to $E(f)$ for some factor of automorphy f of the form

$$f(\lambda, \tilde{x}) = \begin{pmatrix} 1 & \mu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix}.$$

Note that the condition for f to be a factor of automorphy in this case is equivalent to the condition

$$\mu(\lambda + \lambda', \tilde{x}) = \mu(\lambda, \lambda' + \tilde{x}) + \mu(\lambda', \tilde{x}),$$

where we use the additive notation for the group operation since Γ is commutative.

THEOREM 4.5. *f defines the trivial bundle if and only if $\mu(\lambda, \tilde{x}) = \xi(\lambda\tilde{x}) - \xi(\tilde{x})$ for some holomorphic function $\xi : \mathbb{C} \to \mathbb{C}$.*

Proof. We know that E is trivial if and only if $h(\lambda\tilde{x}) = f(\lambda,\tilde{x})h(\tilde{x})$ for some holomorphic function $h : \tilde{X} \to \mathrm{GL}_2(\mathbb{C})$. Let

$$h = \begin{pmatrix} a(\tilde{x}) & b(\tilde{x}) \\ c(\tilde{x}) & d(\tilde{x}) \end{pmatrix},$$

then the last condition is

$$\begin{aligned} \begin{pmatrix} a(\lambda\tilde{x}) & b(\lambda\tilde{x}) \\ c(\lambda\tilde{x}) & d(\lambda\tilde{x}) \end{pmatrix} &= \begin{pmatrix} 1 & \mu(\lambda,\tilde{x}) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a(\tilde{x}) & b(\tilde{x}) \\ c(\tilde{x}) & d(\tilde{x}) \end{pmatrix} \\ &= \begin{pmatrix} a(\tilde{x}) + c(\tilde{x})\mu(\lambda,\tilde{x}) & b(\tilde{x}) + d(\tilde{x})\mu(\lambda,\tilde{x}) \\ c(\tilde{x}) & d(\tilde{x}) \end{pmatrix}. \end{aligned}$$

In particular it means $c(\lambda\tilde{x}) = c(\tilde{x})$ and $d(\lambda\tilde{x}) = d(\tilde{x})$, i.e., c and d are doubly periodic functions on $\tilde{X} = \mathbb{C}$, so they should be constant, i.e., $c(\lambda,\tilde{x}) = c \in \mathbb{C}$, $d(\lambda,\tilde{x}) = d \in \mathbb{C}$.

Now we have

$$\begin{aligned} a(\tilde{x}) + c\mu(\lambda,\tilde{x}) &= a(\lambda\tilde{x}) \\ b(\tilde{x}) + d\mu(\lambda,\tilde{x}) &= b(\lambda\tilde{x}) \end{aligned}$$

which implies

$$\begin{aligned} c\mu(\lambda,\tilde{x}) &= a(\lambda\tilde{x}) - a(\tilde{x}) \\ d\mu(\lambda,\tilde{x}) &= b(\lambda\tilde{x}) - b(\tilde{x}). \end{aligned}$$

Since $\det h(\tilde{x}) \neq 0$ for all $\tilde{x} \in \tilde{X} = \mathbb{C}$ one of the numbers c and d is not equal to zero. Therefore, one concludes that $\mu(\lambda,\tilde{x}) = \xi(\lambda\tilde{x}) - \xi(\tilde{x})$ for some holomorphic function $\xi : \tilde{X} = \mathbb{C} \to \mathbb{C}$.

Now suppose $\mu(\lambda,\tilde{x}) = \xi(\lambda\tilde{x}) - \xi(\tilde{x})$ for some holomorphic function $\xi : \mathbb{C} \to \mathbb{C}$. Clearly for $h(\tilde{x}) = \begin{pmatrix} 1 & \xi(\tilde{x}) \\ 0 & 1 \end{pmatrix}$ one has that $\det h(\tilde{x}) = 1 \neq 0$ and

$$\begin{aligned} f(\lambda,\tilde{x})h(\tilde{x}) &= \begin{pmatrix} 1 & \mu(\lambda,\tilde{x}) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \xi(\tilde{x}) \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \xi(\tilde{x}) + \mu(\lambda,\tilde{x}) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \xi(\lambda\tilde{x}) \\ 0 & 1 \end{pmatrix} = h(\lambda\tilde{x}). \end{aligned}$$

We have shown, that f defines the trivial bundle. This proves the statement of the theorem. □

THEOREM 4.6. *Two factors of automorphy*

$$f(\lambda,\tilde{x}) = \begin{pmatrix} 1 & \mu(\lambda,\tilde{x}) \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad f'(\lambda,\tilde{x}) = \begin{pmatrix} 1 & \nu(\lambda,\tilde{x}) \\ 0 & 1 \end{pmatrix}$$

defining non-trivial bundles are equivalent if and only if

$$\mu(\lambda, \tilde{x}) - k\nu(\lambda, \tilde{x}) = \xi(\lambda\tilde{x}) - \xi(\tilde{x}), \quad k \in \mathbb{C}, \quad k \neq 0$$

for some holomorphic function $\xi : \mathbb{C} = \tilde{X} \to \mathbb{C}$.

Proof. Suppose that the factors of automorphy

$$f(\lambda, \tilde{x}) = \begin{pmatrix} 1 & \mu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix}$$

and

$$f'(\lambda, \tilde{x}) = \begin{pmatrix} 1 & \nu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix}$$

are equivalent. Then there is an equality $f(\lambda, \tilde{x})h(\tilde{x}) = h(\lambda\tilde{x})f(\lambda, \tilde{x})$ for some holomorphic function $h : \mathbb{C} = \tilde{X} \to \mathrm{GL}_2(\mathbb{C})$. Let us write h in the form

$$h(\tilde{x}) = \begin{pmatrix} a(\tilde{x}) & b(\tilde{x}) \\ c(\tilde{x}) & d(\tilde{x}) \end{pmatrix}.$$

Then the condition for equivalence of f and f' can be rewritten as follows:

$$\begin{pmatrix} 1 & \mu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a(\tilde{x}) & b(\tilde{x}) \\ c(\tilde{x}) & d(\tilde{x}) \end{pmatrix} = \begin{pmatrix} a(\lambda\tilde{x}) & b(\lambda\tilde{x}) \\ c(\lambda\tilde{x}) & d(\lambda\tilde{x}) \end{pmatrix} \begin{pmatrix} 1 & \nu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix}$$

After multiplication one obtains

$$\begin{pmatrix} a(\tilde{x}) + c(\tilde{x})\mu(\lambda, \tilde{x}) & b(\tilde{x}) + d(\tilde{x})\mu(\lambda, \tilde{x}) \\ c(\tilde{x}) & d(\tilde{x}) \end{pmatrix} = \begin{pmatrix} a(\lambda\tilde{x}) & a(\lambda\tilde{x})\nu(\lambda, \tilde{x}) + b(\lambda\tilde{x}) \\ c(\lambda\tilde{x}) & c(\lambda\tilde{x})\nu(\lambda, \tilde{x}) + d(\lambda\tilde{x}) \end{pmatrix},$$

which leads to the system of equations

$$\begin{cases} a(\tilde{x}) + c(\tilde{x})\mu(\lambda, \tilde{x}) = a(\lambda\tilde{x}) \\ b(\tilde{x}) + d(\tilde{x})\mu(\lambda, \tilde{x}) = a(\lambda\tilde{x})\nu(\lambda, \tilde{x}) + b(\lambda\tilde{x}) \\ c(\tilde{x}) = c(\lambda\tilde{x}) \\ d(\tilde{x}) = c(\lambda\tilde{x})\nu(\lambda, \tilde{x}) + d(\lambda\tilde{x}). \end{cases}$$

The third equation means that c is a double periodic function. Therefore, c should be a constant function.

If $c \neq 0$ from the first and the last equations using Theorem 4.5 one concludes that f and f' define the trivial bundle.

In the case $c = 0$ one has

$$\begin{cases} a(\tilde{x}) = a(\lambda\tilde{x}) \\ b(\tilde{x}) + d(\tilde{x})\mu(\lambda, \tilde{x}) = a(\lambda\tilde{x})\nu(\lambda, \tilde{x}) + b(\lambda\tilde{x}) \\ d(\tilde{x}) = d(\lambda\tilde{x}), \end{cases}$$

i.e., as above, a and d are constant and both not equal to zero since $\det(h) \neq 0$. Finally one concludes that

$$d\mu(\lambda, \tilde{x}) - a\nu(\lambda, \tilde{x}) = b(\lambda\tilde{x}) - b(\tilde{x}), \quad a, d \in \mathbb{C}, \quad ad \neq 0 \tag{1}$$

Vice versa, if μ and ν satisfy (1) for

$$h(\tilde{x}) = \begin{pmatrix} a & b(\tilde{x}) \\ 0 & d \end{pmatrix}$$

we have

$$\begin{aligned} f(\lambda, \tilde{x})h(\tilde{x}) &= \begin{pmatrix} 1 & \mu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b(\tilde{x}) \\ 0 & d \end{pmatrix} \\ &= \begin{pmatrix} a & b(\tilde{x}) + d\mu(\lambda, \tilde{x}) \\ 0 & d \end{pmatrix} \\ &= \begin{pmatrix} a & b(\lambda\tilde{x}) + a\nu(\lambda, \tilde{x}) \\ 0 & d \end{pmatrix} \\ &= \begin{pmatrix} a & b(\lambda\tilde{x}) \\ 0 & d \end{pmatrix} \begin{pmatrix} 1 & \nu(\lambda, \tilde{x}) \\ 0 & 1 \end{pmatrix} = h(\lambda\tilde{x})f(\lambda, \tilde{x}). \end{aligned}$$

This means that f and f' are equivalent. □

4.2. Higher Dimensional Complex Tori

One can also consider higher dimensional complex tori. Let $\Gamma \subset \mathbb{C}^g$ be a lattice,

$$\Gamma = \Gamma_1 \times \cdots \times \Gamma_g, \quad \Gamma_i = \mathbb{Z} + \mathbb{Z}\tau_i, \quad \operatorname{Im}\tau > 0.$$

Then as for one dimensional complex tori we obtain that $X = \mathbb{C}^g/\Gamma$ is a complex manifold. Clearly the map

$$\mathbb{C}^g \to \mathbb{C}^g/\Gamma = X, \quad x \mapsto [x]$$

is the universal covering of X. Since all vector bundles on $\mathbb{C}^g$ are trivial, we obtain a one-to-one correspondence between equivalence classes of r-dimensional factors of automorphy

$$f : \Gamma \times \mathbb{C}^g \to \mathrm{GL}_r(\mathbb{C})$$

and vector bundles of rank r on X.

Let $\Gamma = \mathbb{Z}^g + \Omega\mathbb{Z}^g$, where Ω is a symmetric complex $g \times g$ matrix with positive definite real part. Note that Ω is a generalization of τ from one dimensional case.

For any theta-characteristic $\xi = \Omega a + b$, where $a \in \mathbb{R}^g$, $b \in \mathbb{R}^g$ there is a holomorphic function $\theta_\xi : \mathbb{C}^g \to \mathbb{C}$ defined by

$$\theta_\xi(z) = \theta_b^a(z, \Omega) = \sum_{n \in \mathbb{Z}^g} \exp(\ \pi i (n+a)^t \Omega (n+a) \tau\ ) \exp(\ 2\pi i (n+a)^t \Omega (z+b)),$$

which satisfies

$$\theta_\xi(\gamma + z) = \exp(2\pi i a^t \gamma - \pi i p^t \Omega p - 2\pi i p^t (z + \xi)) \theta_\xi(z) = e_\xi(\gamma, z) \theta_\xi(z),$$

where $\gamma = \Omega p + q$ and $e_\xi(\gamma, z) = \exp(2\pi i a^t \gamma - \pi i p^t \Omega p - 2\pi i p^t (z + \xi))$. Since

$$e_\xi(\gamma_1 + \gamma_2, z) = e_\xi(\gamma_1, \gamma_2 + z) e_\xi(\gamma_2, z),$$

we conclude that $e_\xi(\gamma, z)$ is a factor of automorphy.

As above $\theta_\xi(z)$ defines a section of $E(e_\xi(\gamma, z))$.

For more detailed information on higher dimensional theta functions see [8, 9, 10].

4.3. Factors of Automorphy depending only on the τ-Direction of the Lattice Γ

Here X is a complex torus, $X = \mathbb{C}/\Gamma$, $\Gamma = \mathbb{Z}\tau + \mathbb{Z}$, $\operatorname{Im} \tau > 0$. Denote $q = e^{2\pi i \tau}$. Consider the canonical projection

$$\mathrm{pr} : \mathbb{C}^* \to \mathbb{C}^*/ < q >, \quad u \to [u] = u < q > .$$

Clearly one can equip $\mathbb{C}^*/ < q >$ with the quotient topology. Therefore, there is a natural complex structure on $\mathbb{C}^*/ < q >$.

Consider the homomorphism

$$\mathbb{C} \overset{\exp}{\to} \mathbb{C}^* \overset{\mathrm{pr}}{\to} \mathbb{C}^*/ < q >, \quad z \mapsto e^{2\pi i z} \mapsto [e^{2\pi i z}].$$

It is clearly surjective. An element $z \in \mathbb{C}$ is in the kernel of this homomorphism if and only if $e^{2\pi i z} = q^k = e^{2\pi i k \tau}$ for some integer k. But this holds if and only if $z - k\tau \in \mathbb{Z}$ or, in other words, if $z \in \Gamma$. Therefore, the kernel of the map is exactly Γ, and we obtain an isomorphism of groups

$$iso : \mathbb{C}/\Gamma \to \mathbb{C}^*/ < q > = \mathbb{C}^*/\mathbb{Z}, \quad [z] \to [e^{2\pi i z}].$$

Since the diagram

$$\begin{tikzcd} \mathbb{C} \arrow[r, "\mathrm{pr}"] \arrow[d, "\exp"'] & \mathbb{C}/\Gamma \arrow[d, leftrightarrow, "iso"] \\ \mathbb{C}^* \arrow[r, "\mathrm{pr}"] & \mathbb{C}^*/\mathbb{Z} \arrow[r, equal] & \mathbb{C}^*/ < q > \end{tikzcd}$$

is commutative, we conclude that the complex structure on $\mathbb{C}^*/<q>$ inherited from $\mathbb{C}/\Gamma$ by the isomorphism *iso* coincides with the natural complex structure on $\mathbb{C}^*/<q>$. Therefore, *iso* is an isomorphism of complex manifolds. Thus complex tori can be represented as $\mathbb{C}^*/<q>$, where $q = e^{2\pi i\tau}$, $\tau \in \mathbb{C}$, $\operatorname{Im}\tau > 0$.

So for any complex torus $X = \mathbb{C}^*/<q>$ we have a natural surjective holomorphic map

$$\mathbb{C}^* \to \mathbb{C}^*/<q>= X, \quad u \to [u].$$

This map is moreover a covering of X. Consider the group $\mathbb{Z}$. It acts holomorphically on $X = \mathbb{C}^*$:

$$\mathbb{Z} \times \mathbb{C}^* \to \mathbb{C}^*, \quad (n, u) \mapsto q^n u.$$

Moreover, since $\operatorname{pr}(q^n u) = \operatorname{pr}(u)$, $\mathbb{Z}$ is naturally identified with a subgroup in the group of deck transformations $\operatorname{Deck}(X/\mathbb{C}^*)$. It is easy to see that $\mathbb{Z}$ satisfies the property (**T**). We obtain that there is a one-to-one correspondence between classes of isomorphism of vector bundles over X and classes of equivalence of factors of automorphy

$$f : \mathbb{Z} \times \mathbb{C} \to \mathrm{GL}_r(\mathbb{C}).$$

Consider the following action of Γ on $\mathbb{C}^*$:

$$\Gamma \times \mathbb{C}^* \to \mathbb{C}^*; \quad (\lambda, u) \mapsto \lambda u =: e^{2\pi i\lambda} u$$

Let $A : \Gamma \times \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$ be a holomorphic function satisfying

$$A(\lambda + \lambda', u) = A(\lambda, \lambda' u) A(\lambda', u) \tag{$*$}$$

for all $\lambda, \lambda' \in \Gamma$. We call such functions $\mathbb{C}^*$-factors of automorphy. Consider the map

$$\mathrm{id}_\Gamma \times \exp : \Gamma \times \mathbb{C} \to \Gamma \times \mathbb{C}^*, \quad (\lambda, x) \to (\lambda, e^{2\pi i x})$$

Then the function

$$f_A = A \circ (\mathrm{id}_\Gamma \times \exp) : \Gamma \times \mathbb{C} \to \mathrm{GL}_r(\mathbb{C})$$

is an r-dimensional factor of automorphy, because

$$\begin{aligned} f_A(\lambda + \lambda', x) = A(\lambda + \lambda', e^{2\pi i x}) &= A(\lambda, e^{2\pi i\lambda'} e^{2\pi i x}) A(\lambda', e^{2\pi i x}) \\ &= A(\lambda, e^{2\pi i(\lambda' + x)}) A(\lambda', e^{2\pi i x}) \\ &= f_A(\lambda, \lambda' + x) f_A(\lambda', x). \end{aligned}$$

So, factors of automorphy on $\mathbb{C}^*$ define factors of automorphy on $\mathbb{C}$.

We restrict ourselves to factors of automorphy $f : \Gamma \times \mathbb{C} \to \mathrm{GL}_r(\mathbb{C})$ with the property

$$f(m\tau + n, x) = f(m\tau, x), \quad m, n \in \mathbb{Z}. \tag{2}$$

It follows from this property that $f(n, x) = f(0, x) = \mathrm{id}_{\mathbb{C}^r}$. Therefore,

$$f(\lambda + k, x) = f(\lambda, k + x)f(k, x) = f(\lambda, k + x) \text{ for all } \lambda \in \Gamma,\ k \in \mathbb{Z}$$

and it is possible to define the function

$$A_f : \Gamma \times \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C}), \quad (\lambda, e^{2\pi i x}) \mapsto f(\lambda, x),$$

which is well-defined because from $e^{2\pi i x_1} = e^{2\pi i x_2}$ follows $x_1 = x_2 + k$ for some $k \in \mathbb{Z}$ and $f(\lambda, x_1) = f(\lambda, x_2 + k) = f(\lambda, x_2)$.

Consider A with the property $A(m\tau + n, u) = A(m\tau, u) =: A(m, u)$. Then clearly $f_A(m\tau + n, u) = f_A(m\tau, u)$. So for any $\mathbb{C}^*$-factor of automorphy $A : \Gamma \times \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$ with the property $A(m\tau + n, u) = A(m\tau, u)$ one obtains the factor of automorphy f_A satisfying (2). We proved the following

THEOREM 4.7. *Factors of automorphy $f : \Gamma \times \mathbb{C} \to \mathrm{GL}_r(\mathbb{C})$ with the property (2) are in a one-to-one correspondence with $\mathbb{C}^*$-factors of automorphy with property $A(m\tau + n, u) = A(m\tau, u)$.*

Now we want to translate the conditions for factors of automorphy with the property (2) to be equivalent in the language of $\mathbb{C}^*$-factors of automorphy with the same property.

THEOREM 4.8. *Let f, f' be r-factors of automorphy with the property (2). Then $f \sim f'$ if and only if there exists a holomorphic function $B : \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$ such that*

$$A_f(m, u)B(u) = B(q^m u)A_{f'}(m, u))$$

for $q := e^{2\pi i \tau}$, where $A(m, u) := A(m\tau, u)$. In this case we also say A_f is equivalent to $A_{f'}$ and write $A_f \sim A_{f'}$.

Proof. Let $f \sim f'$. By definition it means that there exists a holomorphic function $h : \mathbb{C} \to \mathrm{GL}_r(\mathbb{C})$ such that $f(\lambda, x)h(x) = h(\lambda x)f'(\lambda, x)$. Therefore, from $f(n, x)h(x) = h(n + x)f'(n, x)$ and $f(n, x) = f'(n, x) = \mathrm{id}_{\mathbb{C}^r}$ it follows $h(x) = h(n + x)$ for all $n \in \mathbb{Z}$. Therefore, the function

$$B : \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C}), \quad e^{2\pi i x} \mapsto h(x)$$

is well-defined. We have

$$\begin{aligned} A_f(m, e^{2\pi i x})B(e^{2\pi i x}) =& f(m\tau, x)h(x) = h(m\tau + x)f'(m\tau, x) = \\ & B(e^{2\pi i(m\tau + x)})f'(m, e^{2\pi i x}) = B(q^m e^{2\pi i x})A_{f'}(m, e^{2\pi i x}). \end{aligned}$$

Vice versa, let B be such that $A_f(m,u)B(u) = B(q^m A_{f'}(m,u))$. Define $h = B \circ \exp$. We obtain

$$
\begin{aligned}
f(m\tau + n, x)h(x) &= A_f(m\tau + n, e^{2\pi i x})B(e^{2\pi i x}) \\
&= B(q^m e^{2\pi i x})A_{f'}(m\tau + n, e^{2\pi i x}) = B(e^{2\pi i (m\tau + x)})A_{f'}(m\tau + n, e^{2\pi i x}) \\
&= B(e^{2\pi i (m\tau + n + x)})A_{f'}(m\tau + n, e^{2\pi i x}) = h(m\tau + n + x)f'(m\tau + n, x),
\end{aligned}
$$

which means that $f \sim f'$ and completes the proof. □

Remark 4.9. *The last two theorems allow us to embed the set $Z^1(\mathbb{Z}, r)$ of factors of automorphy $\mathbb{Z} \times X \to \mathrm{GL}_r(\mathbb{C})$ to the set $Z^1(\Gamma, r)$. The embedding is*

$$
\Psi : Z^1(\mathbb{Z}, r) \to Z^1(\Gamma, r), \quad f \mapsto g, \quad g(n\tau + m, x) := f(n, x).
$$

Two factors of automorphy from $Z^1(\mathbb{Z}, r)$ are equivalent if and only if their images under Ψ are equivalent in $Z^1(\Gamma, r)$. That is why it is enough to consider only factors of automorphy

$$
\Gamma \times \mathbb{C} \to \mathrm{GL}_r(\mathbb{C})
$$

satisfying (2).

Corollary 4.10. *A factor of automorphy f with property (2) is trivial if and only if $A_f(m,u) = B(q^m u)B(u)^{-1}$ for some holomorphic function $B : \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$.*

Theorem 4.11. *Let A be a $\mathbb{C}^*$-factor of automorphy. $A(m,u)$ is uniquely determined by $A(u) := A(1,u)$.*

$$
A(m,u) = A(q^{m-1}u)\dots A(qu)A(u), \qquad m > 0 \tag{3}
$$

$$
A(-m,u) = A(q^{-m}u)^{-1}\dots A(q^{-1}u)^{-1}, \qquad m > 0. \tag{4}
$$

$A(m,u)$ is equivalent to $A'(m,u)$ if and only if

$$
A(u)B(u) = B(qu)A'(u) \tag{5}
$$

for some holomorphic function $B : \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$. In particular $A(m,u)$ is trivial iff $A(u) = B(qu)B(u)^{-1}$.

Proof. Since $A(1,u) = A(u)$ the first formula holds for $m = 1$. Therefore,

$$
A(m+1,u) = A(1,q^m u)A(m,u) = A(q^m)A(m,u)
$$

and we prove the first formula by induction.

Now $\mathrm{id} = A(0,u) = A(m-m,u) = A(m,q^{-m}u)A(-m,u)$ and hence

$$\begin{aligned} A(-m,u) &= A(m,q^{-m}u)^{-1} = (A(q^{m-1}q^{-m}u)\dots A(qq^{-m}u)A(q^{-m}u))^{-1} \\ &= A(-m,u) = A(q^{-m}u)^{-1}\dots A(q^{-1}u)^{-1} \end{aligned}$$

which proves the second formula.

If $A(m,u) \sim A'(m,u)$ then clearly (5) holds.

Vice versa, suppose $A(u)B(u) = B(qu)A'(u)$. Then

$$\begin{aligned} A(m,u)B(u) &= A(q^{m-1}u)\dots A(qu)A(u)B(u) \\ &= A(q^{m-1}u)\dots A(qu)B(qu)A'(u) \\ &= \dots \\ &= B(q^m u)A'(q^{m-1}u)\dots A'(qu)A'(u) \\ &= B(q^m u)A'(m,u) \end{aligned}$$

for $m > 0$.

Since $A(-m,u) = A(m,q^{-m}u)^{-1}$ we have

$$\begin{aligned} A(-m,u)B(u) &= A(m,q^{-m}u)^{-1}B(u) = (B(u)^{-1}A(m,q^{-m}u))^{-1} \\ &= (B(u)^{-1}A(m,q^{-m}u)B(q^{-m}u)B(q^{-m}u)^{-1})^{-1} \\ &= (B(u)^{-1}B(q^m q^{-m}u)A'(m,q^{-m}u)B(q^{-m}u)^{-1})^{-1} \\ &= (B(u)^{-1}B(u)A'(m,q^{-m}u)B(q^{-m}u)^{-1})^{-1} \\ &= B(q^{-m}u)A'(m,q^{-m}u)^{-1} = B(q^{-m}u)A'(-m,u), \end{aligned}$$

which completes the proof. $\square$

Remark 4.12. *Theorem 4.11 means that all the information about a vector bundle of rank r on a complex torus can be encoded by a holomorphic function $\mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$.*

For a holomorphic function $A : \mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$, let us denote by $E(A)$ the corresponding vector bundle on X.

Theorem 4.13. *Let $A : \mathbb{C}^* \to \mathrm{GL}_n(\mathbb{C})$, $B : \mathbb{C}^* \to \mathrm{GL}_m(\mathbb{C})$ be two holomorphic maps. Then $E(A) \otimes E(B) \simeq E(A \otimes B)$.*

Proof. By theorem 3.5 we have

$$E(A) \otimes E(B) \simeq E(A(n,u)) \otimes E(B(n,u)) \simeq E(A(n,u) \otimes B(n,u)).$$

Since $A(1,u) \otimes B(1,u) = A(u) \otimes B(u)$, we obtain $E(A) \otimes E(B) \simeq E(A \otimes B)$. $\square$

5. Classification of Vector Bundles over a Complex Torus

Here we work with factors of automorphy depending only on τ, i.e., with holomorphic functions $\mathbb{C}^* \to \mathrm{GL}_r(\mathbb{C})$.

5.1. Vector Bundles of Degree Zero

We return to extensions of the type $0 \to I_1 \to E \to I_1 \to 0$, where I_1 denotes the trivial vector bundle of rank 1.

Theorem 4.5 can be rewritten as follows.

Theorem 5.1. *A function*

$$A(u) = \begin{pmatrix} 1 & a(u) \\ 0 & 1 \end{pmatrix}$$

defines the trivial bundle if and only if $a(u) = b(qu) - b(u)$ for some holomorphic function $b : \mathbb{C}^* \to \mathbb{C}$.

Corollary 5.2.

$$A(u) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

defines a non-trivial vector bundle.

Proof. Suppose A defines the trivial bundle. Then $1 = b(qu) - b(u)$ for some holomorphic function $b : \mathbb{C}^* \to \mathbb{C}$. Considering the Laurent series expansion $\sum_{-\infty}^{+\infty} b_k u^k$ of b we obtain $1 = b_0 - b_0 = 0$ which shows that our assumption was false. □

Let $a : \mathbb{C}^* \to \mathbb{C}$ be a holomorphic function such that

$$A_2(u) = \begin{pmatrix} 1 & a(u) \\ 0 & 1 \end{pmatrix}$$

defines non-trivial bundle, i.e., by Theorem 5.1, there is no holomorphic function $b : \mathbb{C}^* \to \mathbb{C}$ such that

$$a(u) = b(qu) - b(u).$$

Let F_2 be the bundle defined by A_2. Then by Theorem 3.3 there exists an exact sequence

$$0 \to I_1 \to F_2 \to I_1 \to 0.$$

For $n \geqslant 3$ we define $A_n : \mathbb{C}^* \to \mathrm{GL}_n(\mathbb{C})$,

$$A_n = \begin{pmatrix} 1 & a & & \\ & \ddots & \ddots & \\ & & 1 & a \\ & & & 1 \end{pmatrix},$$

where empty entries stay for zeros.

Let F_n be the bundle defined by A_n. By (3.3) one sees that A_n defines the extension

$$0 \to I_1 \to F_n \to F_{n-1} \to 0.$$

THEOREM 5.3. *F_n is not the trivial bundle. The extension*

$$0 \to I_1 \to F_n \to F_{n-1} \to 0.$$

is non-trivial for all $n \geqslant 2$.

Proof. Suppose F_n is trivial. Then $A_n(u)B(u) = B(qu)$ for some $B = (b_{ij})_{ij^n}$. In particular it means $b_{ni}(u) = b_{ni}(qu)$ for $i = \overline{1,n}$. Let $b_{ni} = \sum_{-\infty}^{+\infty} b_k^{(ni)} u^k$ be the expansion of b_{ni} in Laurent series. Then $b_{ni}(u) = b_{ni}(qu)$ implies $b_k^{(ni)} = q^k b_k^{(ni)}$ for all k.

Note that $|q| < 1$ because $\tau = \xi + i\eta$, $\eta > 0$ and

$$|q| = |e^{2\pi i\tau}| = |e^{2\pi i(\xi + i\eta)}| = |e^{2\pi i\xi}e^{-2\pi\eta}| = e^{-2\pi\eta} < 1.$$

Therefore, $b_k^{(ni)} = 0$ for $k \neq 0$ and we conclude that b_{ni} should be constant functions.

We also have

$$b_{n-1i}(u) + b_{ni}a(n) = b_{n-1i}(qu).$$

Since at least one of b_{ni} is not equal to zero because of invertibility of B, we obtain

$$a(u) = \frac{1}{b_{ni}}(b_{n-1i}(qu) - b_{n-1i}(u))$$

for some i, which contradicts the choice of a. Therefore, F_n is not trivial.

Assume now, that for some $n > 2$ the extension

$$0 \to I_1 \to F_n \to F_{n-1} \to 0$$

is trivial(for $n = 2$ it is not trivial since F_2 is not a trivial vector bundle). This means

$$A_n \sim \begin{pmatrix} 1 & 0 \\ 0 & A_{n-1} \end{pmatrix},$$

i.e., there exists a holomorphic function $B : \mathbb{C}^* \to \mathrm{GL}_n(\mathbb{C})$, $B = (b_{ij})_{i,j}^n$ such that

$$A_n(u)B(u) = B(qu)\begin{pmatrix} 1 & 0 \\ 0 & A_{n-1} \end{pmatrix}.$$

Considering the elements of the first and second columns we obtain for the first column

$$b_{n1}(u) = b_{n1}(qu),$$

$$b_{i1}(u) + b_{i+11}(u)a(u) = b_{i1}(qu), \quad i < n$$

and for the second column

$$b_{n2}(u) = b_{n2}(qu),$$
$$b_{i2}(u) + b_{i+12}(u)a(u) = b_{i2}(qu), \quad i < n.$$

For the first column as above considering Laurent series we have that b_{n1} should be a constant function. If $b_{n1} \neq 0$ it follows

$$a(u) = \frac{1}{b_{n1}}(b_{n-11}(qu) - b_{n-11}(u)),$$

which contradicts the choice of a. Therefore, $b_{n1} = 0$ and $b_{n-11}(qu) = b_{n-11}(u)$, in other words b_{n-11} is a constant function. Proceeding by induction one obtains that b_{11} is a constant function and $b_{i1} = 0$ for $i > 1$.

For the second column absolutely analogously we obtain a similar result: b_{12} is constant, $b_{i2} = 0$ for $i > 1$. This contradicts the invertibility of $B(u)$ and proves the statement. □

Corollary 5.4. *The vector bundle F_n is the only indecomposable vector bundle of rank n and degree 0 that has non-trivial sections.*

Proof. This follows from [1, Theorem 5]. □

So we have that the vector bundles $F_n = E(A_n)$ are exactly F_n's defined by Atiyah in [1].

Remark 5.5. *Note that constant matrices A and B having the same Jordan normal form are equivalent. This is clear because $A = SBS^{-1}$ for some constant invertible matrix S, which means that A and B are equivalent.*

Consider an upper triangular matrix $B = (b_{ij})_1^n$ of the following type:

$$b_{ii} = 1, \quad b_{ii+1} \neq 0. \tag{6}$$

It is easy to see that this matrix is equivalent to the upper triangular matrix A,

$$a_{ii} = a_{ii+1} = 1, \quad a_{ij} = 0, \quad j \neq i+1, \quad j \neq i. \tag{7}$$

In fact, these matrices have the same characteristic polynomial $(t-1)^n$ and the dimension of the eigenspace corresponding to the eigenvalue 1 is equal to 1 for both matrices. Therefore, A and B have the same Jordan form. By Remark above we obtain that A and B are equivalent. We proved the following:

Lemma 5.6. *A matrix satisfying (6) is equivalent to the matrix defined by (7). Moreover, two matrices of the type (6) are equivalent, i.e., they define two isomorphic vector bundles.*

THEOREM 5.7. *$F_n \simeq S^{n-1}(F_2)$.*

Proof. We know that F_2 is defined by the constant matrix

$$A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We know by Theorem 3.4 that $S^n(F_2)$ is defined by $S^n(A_2)$. We calculate $S^n(f_2)$ for $n \in \mathbb{N}_0$. Since f_2 is a constant matrix, $S^n(f_2)$ is also a constant matrix defining a map $S^n(\mathbb{C}^2) \to S^n(\mathbb{C}^2)$. Let e_1, e_2 be the standard basis of $\mathbb{C}^2$, then $S^n(\mathbb{C})$ has a basis

$$\{e_1^k e_2^{n-k} \mid k = n, n-1, \dots, 0\}.$$

Since $A_2(e_1) = e_1$ and $A_2(e_2) = e_1 + e_2$, we conclude that $e_1^k e_2^{n-k}$ is mapped to

$$\begin{aligned} A_2(e_1)^k A_2(e_1)^{n-k} &= e_1^k (e_1 + e_2)^{n-k} \\ &= e_1^k \sum_{i=0}^{n-k} \binom{n-k}{i} e_1^{n-k-i} e_2^i = \sum_{i=0}^{n-k} \binom{n-k}{i} e_1^{n-i} e_2^i. \end{aligned}$$

Therefore,

$$S^n(A_2) = \begin{pmatrix} 1 & 1 & 1 & \dots & \binom{n}{0} \\ & 1 & 2 & \dots & \binom{n}{1} \\ & & 1 & \dots & \binom{n}{2} \\ & & & \ddots & \vdots \\ & & & & \binom{n}{n} \end{pmatrix},$$

where empty entries stay for zero. In other words, the columns of $S^n(A_2)$ are columns of binomial coefficients. By Lemma 5.6 we conclude that $S^n(A_2)$ is equivalent to A_{n+1}. This proves the statement of the theorem. □

Let E be a 2-dimensional vector bundle over a topological space X. Then there exists an isomorphism

$$S^p(E) \otimes S^q(E) \simeq S^{p+q}(E) \oplus (\det E \otimes S^{p-1}(E) \otimes S^{q-1}(E)).$$

This is the Clebsch-Gordan formula. If $\det E$ is the trivial line bundle, then we have $S^p(E) \otimes S^q(E) \simeq S^{p+q}(E) \oplus S^{p-1}(E) \otimes S^{q-1}(E)$, and by iterating one gets

$$S^p(E) \otimes S^q(E) \simeq S^{p+q}(E) \oplus S^{p+q-2}(E) \oplus \dots \oplus S^{p-q}(E), \quad p \geqslant q. \tag{8}$$

THEOREM 5.8. *$F_p \otimes F_q \simeq F_{p+q-1} \oplus F_{p+q-3} \oplus \dots \oplus F_{p-q+1}$ for $p \geqslant q$.*

Proof. Using Theorem 5.7 and (8) we obtain

$$\begin{aligned} F_p \otimes F_q &\simeq S^{p-1}(F_2) \otimes S^{q-1}(F_2) \\ &\simeq S^{p+q-2}(F_2) \oplus S^{p+q-4}(F_2) \oplus \cdots \oplus S^{p-q}(F_2) \\ &\simeq F_{p+q-1} \oplus F_{p+q-3} \oplus \cdots \oplus F_{p-q+1}. \end{aligned}$$

This completes the proof. □

REMARK 5.9. *The possibility of proving the last theorem using Theorem 5.7 is exactly what Atiyah states in remark 1) after Theorem 9 (see [1, p. 439]).*

We have already given (Corollary 5.4) a description of vector bundles of degree zero with non-trivial sections. We give now a description of all vector bundles of degree zero.

Consider the function $\varphi_0(z) = \exp(-\pi i\tau - 2\pi iz) = q^{-1/2}u^{-1} = \varphi(u)$, where $u = e^{2\pi iz}$. It defines the factor of automorphy

$$e_0(p\tau + q, z) = \exp(-\pi ip^2\tau - 2\pi izp) = q^{-\frac{p^2}{2}}u^{-p}$$

corresponding to the theta-characteristic $\xi = 0$.

THEOREM 5.10. $\deg E(\varphi_0) = 1$, *where as above* $\varphi_0(z) = \exp(-\pi i\tau - 2\pi iz) = q^{-1/2}u^{-1} = \varphi(u)$.

Proof. Follows from Theorem 4.3 for $\xi = 0$. □

THEOREM 5.11. *Let* $L' \in \mathcal{E}(1, d)$. *Then there exists* $x \in X$ *such that* $L' \simeq t_x^* E(\varphi_0) \otimes E(\varphi_0)^{d-1}$.

Proof. Since $E(\varphi_0)^d$ has degree d, we obtain that there exists $\tilde{L} \in \mathcal{E}(1, 0)$ such that $L' \simeq E(\varphi_0)^d \otimes \tilde{L}$. We also know that $\tilde{L} \simeq t_x^* E(\varphi_0) \otimes E(\varphi_0)^{-1}$ (cf. proof of Theorem 4.3 and Theorem 4.4) for some $x \in X$. Combining these one obtains

$$L' \simeq E(\varphi_0)^d \otimes t_x^* E(\varphi_0) \otimes E(\varphi_0)^{-1} \simeq t_x^* E(\varphi_0) \otimes E(\varphi_0)^{d-1}.$$

This proves the required statement. □

THEOREM 5.12. *The map*

$$\mathbb{C}^* / < q > \to \mathrm{Pic}^0(X), \quad \bar{a} \mapsto E(a).$$

is well-defined and is an isomorphism of groups.

Proof. Let $\varphi_0(z) = \exp(-\pi i\tau - 2\pi iz)$ as above. For $x \in X$ consider $t_x^*E(\varphi_0)$, where the map

$$t_x : X \to X, \quad y \mapsto y + x$$

is the translation by x. Let $\xi \in \mathbb{C}$ be a representative of x. Clearly, $t_x^*E(\varphi_0)$ is defined by

$$\varphi_{0\xi}(z) = t_\xi\varphi_0(z) = \varphi_0(z+\xi) = \exp(-\pi i\tau - 2\pi iz - 2\pi i\xi) = \varphi_0(z)\exp(-2\pi i\xi).$$

(Note that if η is another representative of x, then $\varphi_{0\xi}$ and $\varphi_{0\eta}$ are equivalent.) Therefore, the bundle $t_x^*E(\varphi_0) \otimes E(\varphi_0)^{-1}$ is defined by

$$(\varphi_{0\xi}{\varphi_0}^{-1})(z) = \varphi_0(z)\exp(-2\pi i\xi){\varphi_0}^{-1}(z) = \exp(-2\pi i\xi).$$

Since for any $L \in \mathcal{E}(1,0)$ there exists $x \in X$ such that $L \simeq t_x^*E(\varphi_0) \otimes E(\varphi_0)^{-1}$, we obtain $L \simeq E(a)$ for $a = \exp(-2\pi i\xi) \in \mathbb{C}^*$, where $\xi \in \mathbb{C}$ is a representative of x. We proved that any line bundle of degree zero is defined by a constant function $a \in \mathbb{C}^*$.

Vice versa, let $L = E(a)$ for $a \in \mathbb{C}^*$. Clearly, there exists $\xi \in \mathbb{C}$ such that $a = \exp(-2\pi i\xi)$. Therefore,

$$L \simeq E(a) \simeq L(\varphi_{0\xi}{\varphi_0}^{-1}) \simeq t_x^*E(\varphi_0) \otimes E(\varphi_0)^{-1},$$

where x is the class of ξ in X, which implies that $E(a)$ has degree zero. So we obtained that the line bundles of degree zero are exactly the line bundles defined by constant functions.

We have the map

$$\phi : \mathbb{C}^* \to \mathrm{Pic}^0(X), \quad a \mapsto E(a),$$

which is surjective. By Theorem 4.13 it is moreover a homomorphism of groups. We are looking now for the kernel of this map.

Suppose $E(a)$ is a trivial bundle. Then there exists a holomorphic function $f : \mathbb{C}^* \to \mathbb{C}^*$ such that $f(qu) = af(u)$. Let $f = \sum f_\nu a^\nu$ be the Laurent series expansion of f. Then from $f(qu) = af(u)$ one obtains

$$af_\nu = f_\nu q^\nu \text{ for all } \nu \in \mathbb{Z}.$$

Therefore, $f_\nu(a - q^\nu) = 0$ for all $\nu \in \mathbb{Z}$.

Since $f \not\equiv 0$, we obtain that there exists $\nu \in \mathbb{Z}$ with $f_\nu \neq 0$. Hence $a = q^\nu$ for some $\nu \in \mathbb{Z}$.

Vice versa, if $a = q^\nu$, for $f(u) = u^\nu$ we get

$$f(qu) = q^\nu u^\nu = af(u).$$

This means that $E(a)$ is the trivial bundle, which proves $\operatorname{Ker}\phi =< q >$. We obtain the required isomorphism

$$\mathbb{C}^*/ < q > \to \operatorname{Pic}^0(X), \quad \bar{a} \mapsto E(a).$$

This completes the proof. □

THEOREM 5.13. *For any $F \in \mathcal{E}(r,0)$ there exists a unique $\bar{a} \in \mathbb{C}^*/ < q >$ such that $F \simeq E(A_r(a))$, where*

$$A_r(a) = \begin{pmatrix} a & 1 & & \\ & \ddots & \ddots & \\ & & a & 1 \\ & & & a \end{pmatrix}.$$

Proof. By [1, Theorem 5] $F \simeq F_r \otimes L$ for a unique $L \in \mathcal{E}(1,0)$. Since $F_r \simeq E(A_r)$ and $L \simeq E(a)$ for a unique $\bar{a} \in \mathbb{C}^*/ < q >$ we get $F \simeq E(A_r \otimes a)$. So F is defined by the matrix

$$\begin{pmatrix} a & a & & \\ & \ddots & \ddots & \\ & & a & a \\ & & & a \end{pmatrix},$$

where empty entries stay for zeros. It is easy to see that the Jordan normal form of this matrix is

$$\begin{pmatrix} a & 1 & & \\ & \ddots & \ddots & \\ & & a & 1 \\ & & & a \end{pmatrix}.$$

This proves the statement of the theorem. □

5.2. Vector Bundles of Arbitrary Degree

Denote by $E_\tau = \mathbb{C}/\Gamma_\tau$, where $\Gamma_\tau = \mathbb{Z}\tau + \mathbb{Z}$. Consider the r-covering

$$\pi_r : E_{r\tau} \to E_\tau, \quad [x] \mapsto [x].$$

THEOREM 5.14. *Let F be a vector bundle of rank n on E_τ defined by $A(u) = A(1,u) = A(\tau,u)$. Then $\pi_r^*(F)$ is defined by*

$$\tilde{A}(r\tau,u) = \tilde{A}(u) = \tilde{A}(1,u) := A(r\tau,u) = A(q^{r-1}u)\dots A(qu)A(u).$$

Proof. Consider the following commutative diagram.

$$\begin{tikzcd} & \mathbb{C} \arrow[dl, "p_{r\tau}"'] \arrow[dr, "p_\tau"] & \\ E_{r\tau} \arrow[rr, "\pi_r"] & & E_\tau \end{tikzcd}$$

Consider the map

$$E(\tilde{A}) \to \pi_r^*(E(A)) = E_{r\tau} \times_{E_\tau} E(A) = \{([z]_{r\tau},\ [z,v]_\tau) \in E_{r\tau} \times E(A)\},$$
$$[z,v]_{r\tau} \mapsto ([z]_{r\tau},\ [z,v]_\tau).$$

It is clearly bijective. It remains to prove that it is biholomorphic. From the construction of $E(A)$ and $E(\tilde{A})$ it follows that the diagram

$$\begin{tikzcd} E(\tilde{A}) \arrow[dr] & & \\ & \pi_r^*(E(A)) \arrow[r] \arrow[d] & E(A) \arrow[d] \\ & E_{r\tau} \arrow[r] & E_\tau \end{tikzcd}$$

locally looks as

$$\begin{tikzcd} U \times \mathbb{C}^n \arrow[dr] & & \\ & \Delta(U \times U) \times \mathbb{C}^n \arrow[r] \arrow[d] & U \times \mathbb{C}^n \arrow[d] \\ & U \arrow[r, equal] & U \end{tikzcd}, \qquad \begin{tikzcd} (z,v) \arrow[dr, mapsto] & & \\ & ((z,z),v) \arrow[r, mapsto] \arrow[d, mapsto] & (z,v) \arrow[d, mapsto] \\ & z \arrow[r, equal] & z \end{tikzcd},$$

where $\Delta(U \times U)$ denotes the diagonal of $U \times U$.

This proves the required statement. $\square$

THEOREM 5.15. *Let F be a vector bundle of rank n on $E_{r\tau}$ defined by $\tilde{A}(u) = \tilde{A}(r\tau, u)$. Then $\pi_{r*}(F)$ is defined by*

$$A(u) = \begin{pmatrix} 0 & I_{(r-1)n} \\ \tilde{A}(u) & 0 \end{pmatrix}.$$

Proof. Consider the following commutative diagram.

$$\begin{tikzcd} & \mathbb{C} \arrow[dl, "p_{r\tau}"'] \arrow[dr, "p_\tau"] & \\ E_{r\tau} \arrow[rr, "\pi_r"] & & E_\tau \end{tikzcd}$$

Let $z \in \mathbb{C}$. Consider $y = p_{r\tau}(z) \in E_{r\tau}$ and $x = p_\tau(z) = \pi_r p_{r\tau}(z) \in E_\tau$.

Choose a point $b \in \mathbb{C}$ such that $z \in V_b$, where V_b is the standard parallelogram at point b. Clearly $x \in U_b = p_r(V_b)$ and we have the isomorphism $\varphi_b : U_b \to V_b$ with $\varphi_b(x) = z$.

Consider $\pi_r^{-1}(U_b) = W_b \bigsqcup \cdots \bigsqcup W_{b+(r-1)\tau}$, where $y \in W_b$ and $\pi_r|_{W_{b+i\tau}} : W_{b+i\tau} \to U_b$ is an isomorphism for each $0 \leqslant i < r$.

We have

$$\begin{aligned}\pi_{r*}(\mathcal{E}(\tilde{A}))(U_b) &= \mathcal{E}(\tilde{A})(\pi_r^{-1}(U_b)) = \mathcal{E}(\tilde{A})\left(W_b \bigsqcup \cdots \bigsqcup W_{b+(r-1)\tau}\right) \\ &= \mathcal{E}(\tilde{A})(W_b) \oplus \cdots \oplus \mathcal{E}(\tilde{A})(W_{b+(r-1)\tau}),\end{aligned}$$

where $\mathcal{E}(\tilde{A})$ is the sheaf of sections of $E(\tilde{A})$.

Choose $a \in \mathbb{C}$ such that $z \notin V_a$, $z \in V_{a+\tau}$. We have $\varphi_a(x) = z + \tau$. As above, $\pi_r^{-1}(U_a) = W_a \bigsqcup \cdots \bigsqcup W_{a+(r-1)\tau}$ and

$$\begin{aligned}\pi_{r*}(\mathcal{E}(\tilde{A}))(U_a) &=; \mathcal{E}(\tilde{A})(\pi_r^{-1}(U_a)) = \mathcal{E}(\tilde{A})\left(W_a \bigsqcup \cdots \bigsqcup W_{a+(r-1)\tau}\right) \\ &= \mathcal{E}(\tilde{A})(W_a) \oplus \cdots \oplus \mathcal{E}(\tilde{A})(W_{a+(r-1)\tau}).\end{aligned}$$

Since $g_{ab}(x) = A(\varphi_a(x) - \varphi_b(x), \varphi_b(x))$, we obtain

$$g_{ab}(x) = A(\varphi_a(x) - \varphi_b(x), \varphi_b(x)) = A(z + \tau - z, z) = A(\tau, z).$$

Therefore, to obtain $A(\tau, z)$ it is enough to compute $g_{ab}(x)$.

Note that $\pi_{r*}(\mathcal{E}(\tilde{A}))_x = \mathcal{E}(\tilde{A})_y \oplus \cdots \oplus \mathcal{E}(\tilde{A})_{y+(r-1)\tau}$. Note also that g_{ab} is a map from

$$\pi_{r*}(\mathcal{E}(\tilde{A}))(U_b) = \mathcal{E}(\tilde{A})(W_b) \oplus \cdots \oplus \mathcal{E}(\tilde{A})(W_{b+(r-1)\tau})$$

to

$$\pi_{r*}(\mathcal{E}(\tilde{A}))(U_a) = \mathcal{E}(\tilde{A})(W_a) \oplus \cdots \oplus \mathcal{E}(\tilde{A})(W_{a+(r-1)\tau}).$$

One easily sees that $y \in W_b$, $y \in W_{a+(r-1)\tau}$ and $y + i\tau \in W_{b+i\tau}$, $y + i\tau \in W_{a+(i-1)\tau}$ for $0 < i < r$. Therefore, $g_{ab}(x)$ equals

$$\begin{pmatrix} 0 & \tilde{g}_{a\ b+\tau}(y+\tau) & & \\ \vdots & & \ddots & \\ 0 & & & \tilde{g}_{a+(r-2)\tau\ b+(r-a)\tau}(y+(r-1)\tau) \\ \tilde{g}_{a+(r-1)\tau\ b}(y) & 0 & \dots & 0 \end{pmatrix}.$$

It remains to compute the entries of this matrix. Since

$$\begin{aligned}\tilde{g}_{a+(r-1)\tau\ b}(y) &= \tilde{A}(\tilde{\varphi}_{a+(r-1)\tau}(y) - \tilde{\varphi}_b(y), \tilde{\varphi}_b(y)) \\ &= \tilde{A}(z + r\tau - z, z) = \tilde{A}(r\tau, z)\end{aligned}$$

and

$$\begin{aligned}\tilde{g}_{a+(i-1)\tau\ b+i\tau}(y+i\tau) &= \tilde{A}(\tilde{\varphi}_{a+(i-1)\tau}(y+i\tau)-\tilde{\varphi}_{b+i\tau}(y+i\tau),\tilde{\varphi}_{b+i\tau}(y+i\tau)) \\ &= \tilde{A}(z+i\tau-(z+i\tau)=\tilde{A}(0,z+i\tau)=I_n,\end{aligned}$$

one obtains

$$g_{ab}(x)=\begin{pmatrix} 0 & I_n & & \\ \vdots & & \ddots & \\ 0 & & & I_n \\ \tilde{A}(z) & 0 & \dots & 0 \end{pmatrix}.$$

Therefore,

$$A(z)=\begin{pmatrix} 0 & I_n & & \\ \vdots & & \ddots & \\ 0 & & & I_n \\ \tilde{A}(z) & 0 & \dots & 0 \end{pmatrix}=\begin{pmatrix} 0 & I_{(r-1)n} \\ \tilde{A}(u) & 0 \end{pmatrix},$$

which proves the required statement. □

LEMMA 5.16. *Let $A_i \in \mathrm{GL}_n(\mathbb{R})$, $i=1,\dots,n$. Then*

$$\prod_{i=1}^{r}\begin{pmatrix} 0 & I_{(r-1)n} \\ A_i & 0 \end{pmatrix}=\mathrm{diag}(A_r,\dots,A_1)$$

Proof. Straightforward calculation. □

From Theorem 5.14 and Theorem 5.15 one obtains the following:

COROLLARY 5.17. *Let $E(A)$ be a vector bundle of rank n on $E_{r\tau}$, where $A : \mathbb{C}^* \to \mathrm{GL}_n(\mathbb{C}V)$ is a holomorphic function. Then $\pi_r^*\pi_{r*}E(A)$ is defined by*

$$\mathrm{diag}(A(q^{r-1}u),\dots,A(qu),A(u)).$$

In other words $\pi_r^*\pi_{r*}E(A)$ is isomorphic to the direct sum

$$\bigoplus_{i=0}^{r-1}E(A(q^iu)).$$

Proof. We know that $\pi_r^*\pi_{r*}E(A)$ is defined by $B(r,u)$, where

$$B(1,u)=\begin{pmatrix} 0 & I_{(r-1)n} \\ A & 0 \end{pmatrix}.$$

Therefore, using Lemma 5.16, one obtains

$$\begin{aligned} B(r,u) &= \begin{pmatrix} 0 & I_{(r-1)n} \\ A(q^{r-1}u) & 0 \end{pmatrix} \cdots \begin{pmatrix} 0 & I_{(r-1)n} \\ A(qu) & 0 \end{pmatrix} \begin{pmatrix} 0 & I_{(r-1)n} \\ A(u) & 0 \end{pmatrix} \\ &= \operatorname{diag}(A(q^{r-1}u), \dots, A(qu), A(u)), \end{aligned}$$

which completes the proof. □

Corollary 5.18. *Let* $L \in \mathcal{E}(r,0)$*, then* $\pi_r^* \pi_{r*} L = \bigoplus_1^r L$.

Proof. Clear, since $L = E(A)$ for a constant matrix A by Theorem 5.13. □

Note that for a covering $\pi_r : E_{r\tau} \to E_\tau$ the group of deck transformations $\operatorname{Deck}(E_{r\tau}/E_\tau)$ can be identified with the kernel $\operatorname{Ker}(\pi_r)$. But $\operatorname{Ker}\pi_r$ is cyclic and equals $\{1, [q], \dots [q]^{r-1}\}$, where $[q]$ is a class of $q = e^{2\pi i \tau}$ in $E_{r\tau}$. Clearly

$$[q]^*(E(A(u))) = E(A(qu)).$$

Therefore, we get one more corollary.

Corollary 5.19. *Let* ϵ *be a generator of* $\operatorname{Deck}(E_{r\tau}/E_\tau)$*. Then for a vector bundle* E *on* $E_{r\tau}$ *we have*

$$\pi_r^* \pi_{r*} E = E \oplus \epsilon^* E \oplus \cdots \oplus (\epsilon^{r-1})^* E.$$

To proceed we need the following result.

Theorem 5.20 (Oda, [11, Theorem 1.2, (i)]). *Let* $\varphi : Y \to X$ *be an isogeny of* g*-dimensional abelian varieties over a field* k*, and let* L *be a line bundle on* Y *such that the restriction of the map*

$$\Lambda(L) : Y \to \operatorname{Pic}^0(Y), \quad y \mapsto t_y^* L \otimes L^{-1},$$

to the kernel of φ *is an isomorphism. Then* $\operatorname{End}(\varphi_* L) = k$ *and* $\varphi_* L$ *is an indecomposable vector bundle on* X.

Theorem 5.21. *Let* $L \in \mathcal{E}(1,d)$ *and let* $(r,d) = 1$*. Then* $\pi_{r*}(L) \in \mathcal{E}(r,d)$.

Proof. It is clear that $\pi_{r*}L$ has rank r and degree d. It remains to prove that $\pi_{r*}L$ is indecomposable.

We have the isogeny $\pi_r : E_{r\tau} \to E_r$. Since $Y = E_{r\tau}$ is a complex torus (elliptic curve), $Y \simeq \operatorname{Pic}^0(Y)$ with the identification $y \leftrightarrow t_y^* E(\varphi_0) \otimes E(\varphi_0)^{-1}$.

We know that $L = E(\varphi_0)^d \otimes \tilde{L}$ for some $\tilde{L} = E(a) \in \mathcal{E}(1,0)$, $a \in \mathbb{C}^*$. Since $t_y^*(\tilde{L}) = t_y^*(E(a)) = E(a) = \tilde{L}$, as in the proof of Theorem 5.12 one gets

$$
\begin{aligned}
\Lambda(L)(y) &= t_y^*(L) \otimes L^{-1} = t_y^*(E(\varphi_0)^d \otimes \tilde{L}) \otimes (E(\varphi_0)^d \otimes \tilde{L})^{-1} \\
&= t_y^*(E(\varphi_0)^d) \otimes t_y^*(\tilde{L}) \otimes E(\varphi_0)^{-d} \otimes \tilde{L}^{-1} = t_y^*(E(\varphi_0)^d) \otimes E(\varphi_0)^{-d} \\
&= t_y^*(E(\varphi_0^d)(z)) \otimes E(\varphi_0^{-d}) = E(\varphi_0^d(z+\eta)) \otimes E(\varphi_0^{-d}) \\
&= E(\varphi_0^d(z+\eta)\varphi_0^{-d}(z)) = E(\exp(-2\pi i d\eta)) = t_{dy}^*(E(\varphi_0)) \otimes E(\varphi_0)^{-1},
\end{aligned}
$$

where $\eta \in \mathbb{C}$ is a representative of y. This means that the map $\Lambda(L)$ corresponds to the map

$$d_Y : E_{r\tau} \to E_{r\tau}, \quad y \mapsto dy.$$

Since $\operatorname{Ker} \pi_r$ is isomorphic to $\mathbb{Z}/r\mathbb{Z}$, we conclude that the restriction of d_Y to $\operatorname{Ker} \pi_r$ is an isomorphism if and only if $(r,d) = 1$. Therefore, using Theorem 5.20, we obtain the required statement. $\square$

Now we are able to prove the following main theorem:

THEOREM 5.22.

(i) Every indecomposable vector bundle $F \in \mathcal{E}_{E_\tau}(r,d)$ is of the form $\pi_{r'*}(L' \otimes F_h)$, where $(r,d) = h$, $r = r'h$, $d = d'h$, $L' \in \mathcal{E}_{E_{r'\tau}}(1,d')$.

(ii) Every vector bundle of the form $\pi_{r'*}(L' \otimes F_h)$, where L' and r' are as above, is an element of $\mathcal{E}_{E_\tau}(r,d)$.

Proof[1].

(i) By [1, Lemma 26] we obtain $F \simeq E_A(r,d) \otimes L$ for some line bundle $L \in \mathcal{E}(1,0)$. By [1, Lemma 24] we have $E_A(r,d) \simeq E_A(r',d') \otimes F_h$, hence $F \simeq E_A(r',d') \otimes F_h \otimes L$.

Consider any line bundle $\tilde{L} \in \mathcal{E}_{E_{r'\tau}}(1,d')$. Since by Theorem 5.21 $\pi_{r'*}(\tilde{L}) \in \mathcal{E}(r',d')$, it follows from [1, Lemma 26] that there exists a line bundle L'' such that $E_A(r',d') \otimes L \simeq \pi_{r'*}(\tilde{L}) \otimes L''$.

Using the projection formula, we get

$$
\begin{aligned}
F &\simeq \pi_{r'*}(\tilde{L}) \otimes L'' \otimes F_h \\
&\simeq \pi_{r'*}(\tilde{L} \otimes \pi_{r'}^*(L'') \otimes \pi_{r'}^*(F_h)) \\
&\simeq \pi_{r'*}(L' \otimes \pi_{r'}^*(F_h))
\end{aligned}
$$

for $L' = \tilde{L} \otimes \pi_{r'}^*(L'')$.

Since F_h is defined by a constant matrix we obtain by Theorem 5.14 that $\pi_{r'}^*(F_h)$ is defined by $f_h^{r'}$, which is has the same Jordan normal form as f_h. Therefore, $\pi_{r'}^*(F_h) \simeq F_h$ and finally one gets $F \simeq \pi_{r'*}(L' \otimes F_h)$.

(ii) Consider $F = \pi_{r'*}(L' \otimes F_h)$. As above $F_h = \pi_{r'}^*(F_h)$. Using the projection formula we get

$$F = \pi_{r'*}(L' \otimes F_h) = \pi_{r'*}(L' \otimes \pi_{r'}^*(F_h)) = \pi_{r'*}(L') \otimes F_h.$$

By Theorem 5.21 $\pi_{r'*}(L')$ is an element from $\mathcal{E}_{E_\tau}(r', d')$. Therefore, $\pi_{r'*}(L') = E_A(r', d') \otimes L$ for some line bundle $L \in \mathcal{E}_{E_\tau}(1, 0)$. Finally we obtain

$$\begin{aligned} F &= \pi_{r'*}(L') \otimes F_h \\ &= E_A(r', d') \otimes L \otimes F_h \\ &= E_A(r'h, d'h) \otimes L \\ &= E_A(r, d) \otimes L, \end{aligned}$$

which means that F is an element of $\mathcal{E}_{E_\tau}(r, d)$. □

Remark 5.23. *Since any line bundle of degree d' is of the form $t_x^* E(\varphi_0) \otimes E(\varphi_0)^{d'-1}$, Theorem 5.22(i) takes exactly the form of Proposition 1 from [12], which was given without any proof.*

Any line bundle of degree d' over $E_{r\tau}$ is of the form $E(a) \otimes E(\varphi^{d'})$, where $a \in \mathbb{C}^*$. Therefore, $L' \otimes F_h = E(a) \otimes E(\varphi_0^{d'}) \otimes E(A_h) = E(\varphi_0^{d'} A_h(a))$. Using Theorem 5.15 we obtain the following:

Theorem 5.24. *Indecomposable vector bundles of rank r and degree d on E_τ are exactly those defined by the matrices*

$$\begin{pmatrix} 0 & I_{(r'-1)h} \\ \varphi_0^{d'} A_h(a) & 0 \end{pmatrix},$$

where $(r, d) = h$, $r' = r/h$, $d' = d/h$, $\varphi_0(u) = q^{-\frac{r}{2}} u^{-1}$, $q = e^{2\pi i \tau}$, $a \in \mathbb{C}^*$, and

$$A_h(a) = \begin{pmatrix} a & 1 & & \\ & \ddots & \ddots & \\ & & a & 1 \\ & & & a \end{pmatrix} \in \mathrm{GL}_h(\mathbb{C}).$$

Note that if $d = 0$, we get $h = r$, $r' = 1$, and $d' = 0$. In this case the statement of Theorem 5.24 is exactly Theorem 5.13.

References

[1] M.F. Atiyah, *Vector bundles over an elliptic curve*, Proc. Lond. Math. Soc. III. Ser. **7** (1957), 414–452.

[2] C. Birkenhake and H. Lange, *Complex abelian varieties*, second augmented edition, Grundlehren der Mathematischen Wissenschaften 302, Springer, Berlin (2004).
[3] I. Burban and B. Kreussler, *Vector bundles on degenerations of elliptic curves and Yang-Baxter equations*, to appear in Mem. Am. Math. Soc., arXiv:0708.1685v2.
[4] A.L. Carey, K.C. Hannabuss, L.J. Mason and M.A. Singer, *The Landau-Lifshitz equation, elliptic curves and the Ward transform*, Commun. Math. Phys. **154** (1993), 25–47.
[5] O. Forster, *Riemannsche Flächen*, Springer, Berlin (1977).
[6] P. Griffiths and J. Harris, *Principles of algebraic geometry*, John Wiley & Sons, U.S.A. (1978).
[7] O. Iena, *Vector bundles on elliptic curves and factors of automorphy*, Diplomarbeit, Technische Universität Kaiserslautern, Germany (2005).
[8] D. Mumford, *Tata lectures on theta. I: Introduction and motivation: Theta functions in one variable. Basic results on theta functions in several variables. With the assistance of C. Musili, M. Nori, E. Previato, and M. Stillman*, Progress in Mathematics volume 28, Birkhäuser, Basel (1983).
[9] D. Mumford, *Tata lectures on theta. II: Jacobian theta functions and differential equations. With the collaboration of C. Musili, M. Nori, E. Previato, M. Stillman, and H. Umemura*, Progress in Mathematics volume 43, Birkhäuser, Basel (1984).
[10] D. Mumford, M. Nori and P. Norman, *Tata lectures on theta. III*, Progress in Mathematics volume 97, Birkhäuser, Basel (1991).
[11] T. Oda, *Vector bundles on an elliptic curve*, Nagoya Math. J. **43** (1971), 41–72.
[12] A. Polishchuk and E. Zaslow, *Categorical mirror symmetry: the elliptic curve*, Adv. Theor. Math. Phys. **2** (1998), 443–470.
[13] J.-P. Serre, *Géométrie algébrique et géométrie analytique*, Ann. Inst. Fourier **6** (1956), 1–42.

Author's address:

Oleksandr Iena
University of Luxembourg
Campus Kirchberg Mathematics Research Unit
6, rue Richard Coudenhove-Kalergi, 1359 Luxembourg City
Grand Duchy of Luxembourg
E-mail: `oleksandr.iena@uni.lu`

Received May 13, 2011

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 95–110

Language Sampling for Universal Grammars

Luca Bortolussi and Andrea Sgarro

Abstract. *In this paper we present a sampling algorithm for constrained strings representing the state of parameters of a universal grammar. This sampling algorithm has been used to assess statistical significance of the parametric comparison method, a new syntax-based approach to reconstruct linguistic phylogeny.*

Keywords: Monte-Carlo Sampling, Universal Grammars, Language Phylogeny.
MS Classification 2010: 65C60, 68U99

1. Introduction

Historical linguistics is a discipline studying the evolution of languages in the past, with the ultimate aim of gaining a better understanding of past events of human history. A major technique is the construction of phylogenies of current (and extinct) languages. The main methodologies for this task are based on the so called *classical comparative method* [8], exploiting lexical and phonetic relationships to prove language relatedness and to construct language distances (a step required for algorithmic phylogenetic tree reconstruction [4]). Unfortunately, the evolutionary speed of these kinds of data makes them useless to recover relationships more distant in time than 10,000 years.

A recent approach to circumvent these limitations is the *Parametric Comparison Method* (PCM) [5, 6], using syntactic digital data obtained from parameters of the *Universal Grammar* (UG) [3]. Universal Grammar is a recent theory trying to explain language diversity and acquisition in terms of a *finite* number of (binary) switches encoded in the brain that are fixed during language learning and precisely define the syntax of a language. As parameters are supposed to evolve at a much slower rate than lexicon, PCM compares (some) parametric values of languages and extracts phylogenetic information from them. The digital nature of such data parallels the use of DNA information to reconstruct phylogeny of biological species. Unlike with DNA, however, parameters of the UG are interrelated by a complex network of logical dependencies, a fact making statistical analysis of data much more complicated.

In this paper we tackle the problem of sampling uniformly from the set of strings corresponding to admissible values of parameters of a UG, a necessary step in order to perform statistical analysis.

In our setting, there is a fixed number n of boolean parameters of the universal grammar, which are related among them according to logical dependencies. This means that not all possible assignments of boolean values to parameters are *admissible*. Furthermore, certain parameters may not have a defined value under certain circumstances (i.e. depending on the state of related parameters), hence they may have three values: +, -, and 0 (undefined).

The relationships between parameters are given in terms of propositional formulae, where atoms predicate the value of a single parameter. These formulae can be essentially seen as constraints on the space $\mathcal{A}^n$ of strings of length n on the ternary alphabet $\mathcal{A} = \{+, -, 0\}$, so that the space of admissible languages $\mathcal{L}$ is a proper subset of $\mathcal{A}^n$. Hence, the problem we are facing is sampling uniformly from $\mathcal{L}$.

One simple solution is to use a rejection sampling technique: we generate an element of $\mathcal{A}^n$ uniformly (which can be done easily by sampling each position in the string independently) and then we check whether or not it satisfies the constraints. Unfortunately, this approach is unfeasible, because $\mathcal{L}$ is very small with respect to $\mathcal{A}^n$, hence this method is much too costly. The solution we propose here is to use a modification of this basic rejection sampling technique: we sample from a subspace $\mathcal{L}_0$ such that $\mathcal{L} \subset \mathcal{L}_0 \subset A^n$, and such that the relative dimension of $\mathcal{L}$ in $\mathcal{L}_0$ is reasonably large. The main difficulty is to find such a space $\mathcal{L}_0$ together with an efficient algorithm to sample uniformly from it.

Our main strategy is to investigate the structure of the constraints on the languages. If these constraints are simple enough, we can build a data structure enabling fast sampling. If all constraints were of this kind, we would have an efficient sampler. However, just a small fraction of constraints is simple enough. In order to make the method feasible, we preprocess the space of parameters by merging some of them in order to simplify the structure of the constraints. This approach is fit to deal with the number of parameters typical of the linguistical applications we have to face.

In the rest of the paper, we first present formally the data we have to work with, namely parameters and rules (Section 2). Subsequently, we recall the basics of rejection sampling (Section 3), and then we present our fast sampling based on a data structure managing simple constraints, called *sampling structure* (Section 4). This part is the core of the paper: we define formally the data structure used for the sampling, proving the correctness of our algorithm. In Section 5, we deal with the problem of merging parameters together so as to simplify the structure of rules. Finally, in Section 6, we discuss the performances of the algorithm on real data and draw some conclusions.

2. Parameters and Rules

As customary in science, we work with a model of a certain aspect of reality. Our interest is in the historical evolution of languages, and we focus on a particular description of languages based on the universal grammar [3, 6]. Essentially, each language is characterized by a set of boolean parameters, which fix a certain feature of the syntax of the language, mimicking the way language grammar is learnt and "stored" by human brain. Abstracting from the precise meaning of parameters, each language is represented by a tuple of values for the parameters considered.

However, parameters are not independent of one another, but rather they are connected by an intricate network of logical relationships. More precisely, it is meaningful to consider a parameter only if certain other parameters have specific values. For instance, a parameter P_2 may be considered only if the parameter P_1 has value $+$ (i.e. the characteristic it encodes is present). These relationships build up a causal structure on parameters, which is free of logical vicious circles, meaning that parameters can be ordered so that the meaningfulness of parameter P_j depends only on parameters P_i, for $i < j$. Rules for parameters are usually expressed by propositional formulae, specifying which values of parameters are required in order for parameter P_j to be meaningful. For instance, the formula for parameter P_3 may be $\phi_3 = (x_1 = +) \wedge (x_2 = -)$.

We introduce now some notational conventions that will be used throughout the paper. We assume we have n parameters, indicated by $P_1, \ldots, P_n$. Each parameter P_i can take values in a finite domain $\mathcal{D}_i$. Usually, $\mathcal{D}_i = \{+, -\}$, so that P_i is binary. However, for reasons that will be made clear in Section 5, we prefer to work in a more general setting, without restricting the cardinality of $\mathcal{D}_i$ (essentially, we want to deal with macro-parameters, constructed by merging together several "basic" parameters, thus having more than 2 possible states). In addition, each parameter can take a special value, 0, which indicates that it is not meaningful in the current context (i.e., given the value of the parameters it depends on).

Hence, the proper domain for parameter P_i is $\tilde{\mathcal{D}}_i = \mathcal{D}_i \cup \{0\}$, and the space of tuples we need to consider is $\mathcal{D} = \prod_{i=1}^{n} \tilde{\mathcal{D}}_i$.

Only the elements of $\mathcal{D}$ that satisfy a given set of constraints, expressed in terms of rules, are *admissible*. A *rule* is a pair (P_j, ϕ_j), where P_j is a parameter and ϕ_j is a propositional formula. The atoms are equalities of the form $x_i = a_i$, where x_i is a variable referring to the parameter P_i and $a_i \in \mathcal{D}_i$ is a possible value of P_i. We require that all the variables appearing in the formula ϕ_j are among $x_1, \ldots, x_{j-1}$. We consider a set of rules $\mathcal{R}$, containing one rule for each parameter P_j (note that a formula ϕ can be a tautology). We say that a tuple $\mathbf{a} \in \mathcal{D}$ *satisfies* a rule (P_j, ϕ_j) in $\mathcal{R}$ if and only if the formula $\phi_j[\mathbf{a}]$ is true. A tuple $\mathbf{a} \in \mathcal{D}$ is *admissible* if and only if it satisfies all the rules of the set $\mathcal{R}$.

3. Rejection Sampling

Rejection sampling is a standard sampling technique [2, 7] to sample indirectly from a target probability distribution. More precisely, suppose we want to sample from a distribution $q_1(x)$ on space X, but we do not know how to sample from q_1. However, we have a sampling algorithm for another distribution $q_2(x)$ on X, and we know that there exists a constant $M > 0$ such that, for each $x \in X$, $q_1(x) \leq Mq_2(x)$. Rejection sampling consists in sampling from q_2 and accepting a sample with probability $\frac{q_1(x)}{Mq_2(x)}$. More precisely, the sampling algorithm works as follows:

1. sample x from q_2;
2. sample u from the uniform distribution in $[0, 1]$;
3. accept x if $u < \frac{q_1(x)}{Mq_2(x)}$.

In our context, rejection sampling in even simpler. Suppose we want to sample uniformly from a subset $\mathcal{L} \subseteq \mathcal{D}$, but we only have an uniform sampler for an intermediate space $\mathcal{L}_0$, such that $\mathcal{L} \subseteq \mathcal{L}_0 \subseteq \mathcal{D}$. In this setting, an element $x \in \mathcal{L}_0$ is sampled with probability $\frac{1}{|\mathcal{L}_0|}$, and the constant M is $M = \frac{|\mathcal{L}_0|}{|\mathcal{L}|}$. This implies that the acceptance rule simplifies to: *accept x if and only if it belongs to $\mathcal{L}$.*

The problem with rejection sampling lies in the constant M. In fact, M can be seen as the average number of trials one has to do in order to generate an element of $\mathcal{L}$. Hence, if M is very large, the rejection sampling becomes highly inefficient.

In our application context, if we choose $\mathcal{L}_0 = \mathcal{D} = \mathcal{A}^n$, from which uniform sampling is very easy to implement (just select a value for each parameter uniformly and independently), we obtain a constant which is incredibly large (of the order of 10^{18} in the case of [6], see also Section 6). Hence, the only way of using rejection sampling in our context is to identify a much smaller super-space of $\mathcal{L}$. We tackled the problem in the following way:

- we identified a simple form for rules that allows fast uniform sampling, using a dedicated data structure;
- we separated actual rules into two subsets, those that allow fast sampling and the others;
- we sample parameters governed by simple rules using the dedicated sampling and we sample the other parameters uniformly, then accept if all rules are satisfied (i.e. if the language so generated is in $\mathcal{L}$).

In this way, we are sampling from an intermediate space $\mathcal{L}_0$, as we never generate sequences of parameters that violate a simple rule.

Unfortunately, also this approach is limited, because the set of bad rules is too large (resulting in a large constant M). In order to reduce its size (hence the constant M), in Section 5 we merge some parameters into macro-parameters. In this way, we remove some rules (as they are taken into account in defining the allowed values of the macro-parameter) and we simplify other rules, making them simple enough to be dealt with by means of the fast sampler.

4. Sampling Structures

Sampling directly admissible languages according to a uniform probability distribution is complicated because of the complexity of the rules. However, a direct sampling algorithm may be feasible if we impose sufficient restrictions on the rules. In particular, we will consider restrictions allowing the construction of a collection of decision trees to uniformly sample the value of parameters in *linear* time with respect to n.

In order to construct our decision trees, the meaningfulness of a parameter P_j should depend only on another *single* parameter P_i, $i < j$. This surely holds if the formula for P_j contains a single atom, namely it is of the form $x_i = a_i$ (*simple rules*). This also holds for formulae that are disjunctions of atoms, such that only one atom can be true at a given time (*exclusive OR rules*), due to constraints imposed by rules on the parameters involved in the disjunction. In this latter case, we just need to consider one parameter, the one whose atom is true. Hence, in the following we assume that the set of rules $\mathcal{R}$ contains only tautologies, simple rules and exclusive OR rules;[1] in subsection 4.2 we show how to check whether an OR rule is indeed exclusive. A set $\mathcal{R}$ of this kind will be dubbed *simple*. Under this restriction, we will show how to build a set of trees (a forest) and how to use it for sampling uniformly. We will refer to this forest as the *sampling structure* associated with $\mathcal{R}$.

4.1. Definition of the Sampling Structure

The sampling structure associated with a simple set of rules $\mathcal{R}$ contains two kinds of nodes: *parameter nodes*, or *p-nodes*, and *value nodes*, or *v-nodes*. The former represent parameters, while the latter encode the different values that a parameter can take. These nodes will be annotated: p-nodes store the index of the parameter they are associated with, while v-nodes store both the index of the parameter and the value of the element they refer to. Edges in this graph represent two different things: edges from p-nodes to v-nodes simply connect each p-node with all the possible values its parameter can take, while edges from

[1] Note that simple rules are special cases of exclusive OR rules, with just 1 disjunct.

v-nodes to p-nodes encode the dependencies between parameters through rules. Essentially, there will be an edge from a v-node for parameter P_i and value a_i to a p-node for parameter P_j if and only if $x_i = a_i$ is a disjunct in the rule for P_j. Furthermore, we will require that this forest contains the information of all the rules in $\mathcal{R}$. We now provide a formal definition of the sampling structure (a forest, actually) associated with a given simple set of rules $\mathcal{R}$.

DEFINITION 4.1. *A* sampling structure *for a simple set of rules $\mathcal{R}$ is a tuple $T = (P, V, E, \iota, \lambda)$, where:*

- *P is the set of p-nodes;*
- *V is the set of v-nodes;*
- *$E \subset (P \times V) \cup (V \times P)$ is the set of edges;*
- *$\iota : P \cup V \to \{1, \dots, n\}$ associates with each node the index of the parameter it refers to;*
- *$\lambda : V \to \bigcup_{i=1}^{n} \mathcal{D}_i$ associates with each v-node v a value $\lambda(v) \in \mathcal{D}_{\iota(v)}$.*

T satisfies the following properties:

1) *for all $i = 1, \dots, n$, there exists $p \in P$ such that $\iota(p) = i$ (completeness of p-nodes);*
2) *$(p, v) \in E$ implies that $\iota(p) = \iota(v)$ (coherence between a p-node and its children v-node);*
3) *for each $v \in V$ there exists $p \in P$ such that $(p, v) \in E$ (each v-node has a p-parent);*
4) *for each $p \in P$, there exist $v_1, \dots, v_h \in V$, $h = |\mathcal{D}_{\iota(p)}|$, such that $(p, v_i) \in E$ and $\lambda(v_i) \neq \lambda(v_j)$, for each $i \neq j$ (each p-node has children v-nodes for all possible values);*
5) *if $(P_j, (x_{i_1} = a_{i_1}) \vee \dots \vee (x_{i_k} = a_{i_k})) \in \mathcal{R}$, then for each $v \in V$ such that $\iota(v) = i_s$ and $\lambda(v) = a_{i_s}$, there exists $p \in P$ with $\iota(p) = j$ and $(v, p) \in E$ (disjuncts imply edges);*
6) *for each $(v, p) \in E$, $x_{\iota(p)} = \lambda(v_{\iota(v)})$ is a disjunct in the rule for $P_{\iota(p)}$ (edges imply disjuncts).*

In graph-theoretic terms, the previous definition corresponds to a forest, whose roots correspond to independent parameters, i.e. those having a tautology assigned by the rules in $\mathcal{R}$. Each parameter can be associated with one or more p-nodes. Exclusive OR rules with more than one disjunct multiply the

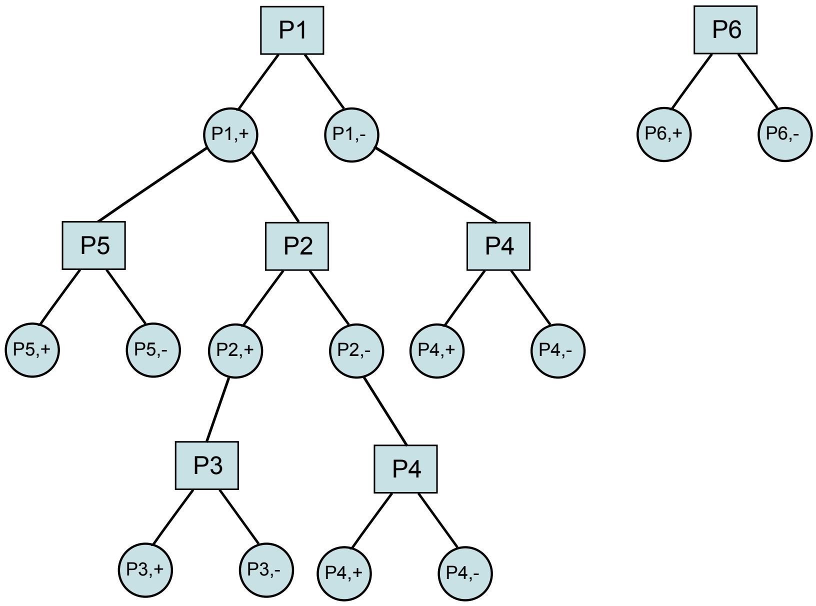


Figure 1: An example of a sampling structure for parameters $P_1, \ldots, P_6$, all taking values in $\{+, -\}$, and subject to rules $\mathcal{R} = \{(P_1, true), (P_2, x_1 = +), (P_3, x_2 = +), (P_4, (x_1 = -) \vee (x_2 = -)), (P_5, x_1 = +), (P_6, true)\}$.

occurrences of p-nodes for a given parameter: a chain of exclusive OR rules can provoke an exponential growth of p-nodes (with respect to n). However, this combinatorial explosion can be readily tamed, see Section 4.5 below.

As an example, consider parameters $P_1, \ldots, P_6$, all taking values in $\{+, -\}$, and subject to the following set of rules $\mathcal{R} = \{(P_1, true), (P_2, x_1 = +), (P_3, x_2 = +), (P_4, (x_1 = -) \vee (x_2 = -)), (P_5, x_1 = +), (P_6, true)\}$. Observe that $\mathcal{R}$ is simple because the rule for P_4 is a disjunctive OR rule: P_2 is defined if and only if P_1 is set to $+$. The sampling structure associated with this set of rules and parameters is shown in Figure 1.

4.2. Exclusive OR Rules and Sampling Tree Structure

Before entering into the details of using a sampling structure to sample admissible languages, we describe some structural properties that can be used to check if the OR rules in the set of rules $\mathcal{R}$ are indeed exclusive. We collect these properties in the following proposition.

Proposition 4.2. *Let $T = (P, V, E, \iota, \lambda)$ be a sampling structure for the set of simple rules $\mathcal{R}$.*

1) *For each independent parameter P_j of $\mathcal{R}$, there is a single p-node $p \in P$ such that $\iota(p) = j$, and it is the root of a tree in the forest T.*

2) *All p-nodes for the same parameter P_j belong to the same tree of the forest T.*

3) *Let p_1, p_2 be two nodes for the same parameter, i.e. $\iota(p_1) = \iota(p_2)$, and let w be the lowest common ancestor of p_1 and p_2, $w = LCA(p_1, p_2)$. Then $w \in P$ (i.e. w is a p-node).*

Proof.

1) This is straightforward, otherwise the property 6 of Definition 4.1 would be violated.

2) Suppose not, and let P_j be the first parameter violating the property, so that there are two nodes p_1 and p_2 for P_j belonging to different trees in the forest. The ancestor p-nodes of p_1 and p_2 are all different (due to the choice of P_j), hence the parent p-nodes and v-nodes of p_1 and p_2 correspond to two different disjuncts in the rule for P_j. Disjointness of ancestor p-nodes implies that there is an admissible tuple that satisfies both disjuncts, constructed by assigning to the parameter of each p-node p the value of its children v-nodes v in the path to p_1 or p_2, i.e. by assigning to $P_{\iota(p)}$ the value $\lambda(v)$. This contradicts the hypothesis that $\mathcal{R}$ is simple.

3) Suppose not, and let P_j be the first parameter violating the property, so that there are two nodes p_1 and p_2 for P_j whose LCA is a v-node v. Due to the choice of P_j, the p-nodes in the paths from v to p_1 and from v to p_2 correspond to disjoint parameters, hence the parent p-nodes and v-nodes of p_1 and p_2 correspond to two different disjuncts of the OR rule for P_j. Reasoning as in the previous point, there is an admissible tuple that satisfies both disjuncts, a contradiction. □

The last proposition gives a way to check if a set of rules containing just OR rules violates the property of being simple (namely, if all OR rules are exclusive). In fact, one just has to construct the sampling structure of $\mathcal{R}$ and check if points 2 and 3 of the previous proposition are violated or not. If they are violated, one can conclude that some OR rules are not exclusive. Actually, one can also prove the inverse of points 2 and 3: if the sampling structure is such that all p-nodes for the same parameter are in the same tree and their

LCA is always a p-node, then the rule set $\mathcal{R}$ is simple.[2] This provides a characterization of simple rule sets in terms of their sampling structures.

4.3. Instances and Languages

The key notion to use a sampling structure T to sample admissible languages is that of an *instance* of T. Intuitively, an instance is a sub-forest of T that can be mapped to a single admissible language. It is constructed starting from the roots and recursively picking only one child for each p-node, and all children of v-nodes. Choosing a child of a p-node corresponds to fixing the value of a parameter. Choosing all children of a v-node is necessary because we need to fix all meaningful parameters, i.e. those whose formulae in $\mathcal{R}$ are true.

We first formally define an instance of T, and then prove that instances of T and admissible languages are in bijection.

DEFINITION 4.3. *Let $T = (P, V, E, \iota, \lambda)$ be a sampling structure for the set of simple rules $\mathcal{R}$. An* instance *I of T is a subgraph $(P_I, V_I, E_I, \iota, \lambda)$ of T such that:*

1) *roots of T are in I;*
2) *if $p \in P_I$ then there exists an unique $v \in V$ such that $v \in V_I$ and $(p, v) \in E_I$ (an instance contains just one child for each p-node);*
3) *if $v \in V_I$, then for each $p \in P$ such that $(v, p) \in E$, it holds that $p \in P_I$ and $(v, p) \in E_I$ (an instance contains all children of v-nodes).*

In order to prove that instances and admissible languages are in one to one correspondence, we will show how we can construct an admissible language from an instance and, viceversa, how to construct an instance from an admissible language.

First, consider an instance I, and define a language $\mathbf{a}^I$ according to the following rule:

$$a_j^I = \begin{cases} \lambda(v), & if\ v \in V_I \wedge \exists p \in P_I : \iota(p) = j \wedge (p, v) \in E_I \\ 0, & \text{otherwise.} \end{cases}$$

In order for $\mathbf{a}^I$ to be well defined, the number of p-nodes for parameter P_j in an instance must be at most one. This is guaranteed by the following lemma.

[2]The proof works as follows. Suppose the OR rule for parameter P_j is not exclusive. Then there is an admissible language that satisfies two disjuncts of this rule. Use the value of parameters in this language to choose a child v-node for each p-node in T, like in the language to instantiate construction of the next section. In this way, we will find two p-nodes for P_j. But their LCA must be a p-node, hence we should have chosen two different v-nodes for it, a contradiction.

LEMMA 4.4. *Let I be an instance of a sampling structure $T = (P, V, E, \iota, \lambda)$ for simple rules $\mathcal{R}$. For each parameter P_j, there exists at most one node $p \in P_I$ such that $\iota(p) = j$.*

Proof. Suppose not, and let P_j be the first parameter having two p-nodes in I, say p_1 and p_2. Then, as each p-node in I has just one child and p_1 and p_2 belong to the same tree of T, the LCA of p_1 and p_2 must be a v-node, in contradiction with point 3 of Proposition 4.2. □

We now need to prove that $\mathbf{a}^I$ satisfies all the rules of $\mathcal{R}$.

LEMMA 4.5. *The tuple $\mathbf{a}^I$ satisfies all the rules in $\mathcal{R}$.*

Proof. We prove the lemma by finite induction on the parameter index j.

($j = 1$) The constraint formula for P_1 is always a tautology, hence there is only one p-node for it, and all its values are admissible.

($j - 1 \Rightarrow j$) We first consider the case in which $a_j^I = 0$. In this case, the boolean formula for P_j must be false, otherwise there would be a true disjunct, say $x_i = a_i^I$, with $i < j$, so that there would be a v-node v_i in I with $\lambda(v_i) = a_i^I$ such that one of its children p-nodes would not belong to I. Now, suppose $a_j^I \neq 0$. Let p be the node for P_j in I, and let v be its parent v-node. Then, by definition of sampling structure $x_{\iota(v)} = \lambda(v) = a_{\iota(v)}^I$ is a disjunct in the boolean formula for P_j, which is therefore true. □

We now consider an admissible language $\mathbf{a}$ and associate to it an instance $I_{\mathbf{a}}$. The construction is done recursively in a simple way:

- add to $I_{\mathbf{a}}$ the p-node for P_1, its child v-node v with $\lambda(v) = a_1$ and all p-nodes children of v;
- if $a_j \neq 0$, consider the only p-node for parameter P_j in the instance $I_{\mathbf{a}}$ being constructed, and add to $I_{\mathbf{a}}$ its child v-node v with $\lambda(v) = a_j$ and all p-nodes children of v
- if $a_j = 0$, do nothing;

We observe that step 2 of the previous construction can always be carried out. In fact, there is always just one p-node for P_j in case $a_j \neq 0$, as there must be a v-node corresponding to one of the disjuncts in the rule for P_j previously inserted in $I_{\mathbf{a}}$ ($\mathbf{a}$ is admissible) and all its children are also in $I_{\mathbf{a}}$. Hence, the following lemma holds.

LEMMA 4.6. *$I_{\mathbf{a}}$ is an instance of T.* □

We have just defined two mappings, one from admissible languages to instances and the other from instances to admissible languages. By construction of these mappings, it holds that they are one the inverse of the other, namely $I_{\mathbf{a}^I} = I$ and $\mathbf{a}^{I_\mathbf{a}} = \mathbf{a}$. This is sufficient to prove the following theorem.

THEOREM 4.7. *Let $\mathcal{R}$ be a simple set of rules and T be its sampling structure. Then instances of T and admissible languages of $\mathcal{R}$ are in bijection.* □

This theorem is the key to the sampling algorithm: in order to sample uniformly admissible languages for a simple rule set, we can sample uniformly instances from the associated sampling structure.

4.4. Uniform Sampling of Instances

We now turn to detail the sampling mechanism of instances, proving that it samples instances according to the uniform probability distribution. A key step towards the sampling algorithm is to count all possible instances of each tree and subtree of a sampling structure. This is necessary because, in order to sample uniformly, we need to know how many instances can be generated choosing one or another child v-node of a p-node. The counting function is defined inductively on the height of nodes of each tree of the sampling structure, distinguishing between v-nodes and p-nodes, as instances treat them differently.

DEFINITION 4.8. *Let $T = (P, V, E, \iota, \lambda)$ be a sampling structure for simple rules $\mathcal{R}$. The instance-counting function $N : P \cup V \to \mathbb{N}$ is defined recursively as follows:*

1) *for each v-node v of height $h(v) = 0$, $N(v) = 1$;*
2) *for each p-node p of height $h > 0$, with children $v_1, \ldots, v_k$, $N(p) = N(v_1) + \ldots + N(v_k)$;*
3) *for each v-node v of height $h > 0$, with children $p_1, \ldots, p_k$, $N(v) = N(p_1) \cdot \ldots \cdot N(p_k)$.*

The correctness of the previous definition is easily proved by induction on the height h of a node. Intuitively, at each internal p-node, we add to an instance just one child, hence we need to add up the number of instances of the children, while for internal v-nodes, we add all children to each instance, hence we need to consider all possible combinations , hence take the product. If $p_1, \ldots, p_k$ are the roots of the distinct trees of T, then the total number of instances (and, accordingly, of admissible languages) is $N(T) = N(p_1) \cdot \ldots \cdot N(p_k)$.

The sampling algorithm needs to pick an instance among the possible ones, according to the uniform distribution. The idea is to choose an instance, following Definition 4.3, by choosing a v-node for each p-node inserted in the instance. More precisely, we say that a p-node is *active* if it has been inserted in the instance, but one of its children v-nodes has still to be selected. The sampling algorithm that constructs $I = (P, V, E)$ is the following:

$P \leftarrow$ {roots of T}
$A \leftarrow$ {roots of T} {A is a set containing active p-nodes}
while $A \neq \emptyset$ **do**
  Remove p from A
  Choose child v_i of p among $v_1, \ldots, v_k$ with probability $\frac{N(v_i)}{N(p)}$
  Add v_i to V and (p, v_i) to E
  Add the children nodes $p_1, \ldots, p_s$ of v_i to P and to A, and (v_i, p_i) to E
end while

The previous algorithm samples an instance with uniform probability. The fact that it generates instances is straightforward (it replicates the recursive definition of an instance). As for the uniform probability, observe that the probability π with which a generic instance is generated is the product of the probabilities of the choices performed in its internal p-nodes. Now, pick a generic factor of this product π, namely $\frac{N(v)}{N(p)}$. If v is an internal v-node, then all its children $p_1, \ldots, p_s$ are inserted in the instance, and they contribute to the product π with factors $\frac{N(v_i)}{N(p_i)}$. Now, the numerator $N(v) = N(p_1) \cdot \ldots \cdot N(p_s)$ cancels out with the terms $N(p_i)$ of the denominator. Therefore, the only factors left at numerator are $N(v)$ for v leaf, hence the numerator equals 1. Similarly, the only factors that remain at the denominator are the terms $N(p)$, for each root p of T. It follows that $\pi = \frac{1}{N(T)}$, showing that instances are sampled uniformly.

4.5. Compact Sampling Trees

The complexity of the sampling algorithm of the previous section is linear in the size of T. However, as anticipated at the end of Section 4, the size of T can grow quicker than linearly with n, due to the fact that exclusive OR rules may introduce many p-nodes for the same parameter P_j. For instance, consider a sampling structure for parameters $P_1, \ldots, P_n$, with values in $\{+, -\}$, and subject to rules $\mathcal{R} = \{(P_1, true), (P_2, x_1 = +), \ldots, (P_i, (x_{i-1} = +) \vee (x_{i-2} = -)), \ldots\}$. It is easy to see that there are exactly $i - 1$ p-nodes for parameter P_i ($i > 2$), hence the total number of p-nodes is quadratic in n (see Figure 2 left).

This combinatorial growth can be avoided by relaxing the constraint that T is a collection of trees, and allowing it to be a collection of direct acyclic graphs (DAGs). The key observation, in fact, is that all subtrees rooted at p-nodes p_i for parameter P_j are isomorphic. Hence, we can merge them into a single tree, effectively inserting one single p-node for P_j. If we perform this collapsing from the last parameter backwards, at the end we obtain a collection of DAGs,

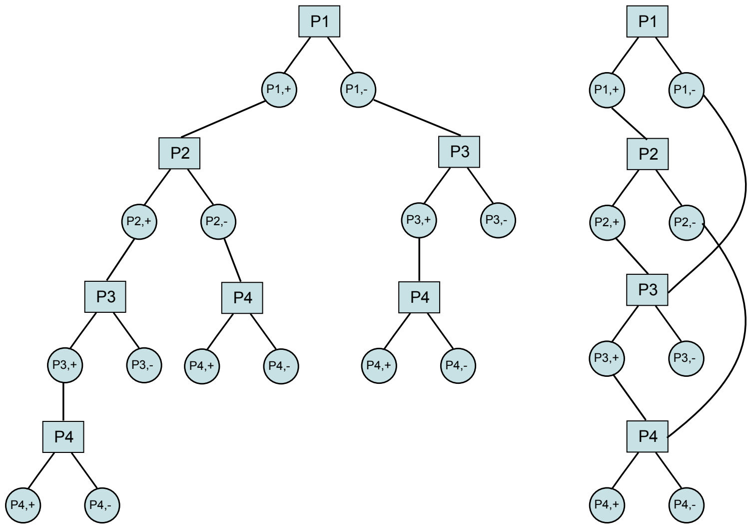


Figure 2: An example of a sampling structure (left) and of a compact sampling structure (right) for parameters $P_1, \ldots, P_4$, all taking values in $\{+,-\}$, and subject to rules $\mathcal{R} = \{(P_1, true), (P_2, x_1 = +), (P_3, (x_1 = +) \vee (x_2 = -)), (P_4, (x_2 = +) \vee (x_3 = -))\}$.

with exactly n distinct p-nodes, one for each parameter. We call this object a *compact sampling structure* (see Figure 2 right).

Instances for a compact sampling structure can be defined similarly to sampling structures. It is easy to see that the set of instances is the same for both structures (the exclusive nature of OR rules implies that we can reach a p-node only through a single incoming edge at a time). Hence, sampling structures and compact sampling structures can be used interchangeably. This linear representation allows one to sample an admissible language in time $O(n)$, using $O(n + |\bigcup_i \mathcal{D}_i|)$ extra space.

5. Parameter Merging

In the previous section, we discussed how to construct an efficient data structure to sample uniformly from a space of languages defined by a restricted set of rules, namely exclusive disjunctions of literals. Unfortunately, the rules governing real parameters of the universal grammar can be dramatically more complex. In these cases, the previous approach is not applicable, hence we have to resort to the rejection sampling strategy discusses in Section 3. However, this still results in a sample space which is too large in practice. In order to reduce the dimension of the sampling space overapproximating the space of admissible languages, we can use a different approach, merging sets of parameters into a bigger macro-parameter so that the resulting set of rules simplifies. Therefore, the parameters to be merged are those involved in formulae that are not exclusive disjunctions.

As an example, consider four parameters, P_1, P_2, P_3, P_4, taking values in $\{+, -\}$ and subject to rules $\mathcal{R} = \{(P_1, true), (P_2, x_1 = +), (P_3, true), (P_4, (x_2 = +) \wedge (x_3 = -))\}$. The rules for P_1, P_2, and P_3 are simple, but the rule for P_4 is not. A possible solution is merging parameters P_2 and P_3 into a macro-parameter $P_{2,3}$, which can take four values, i.e. $\{++, +-, -+, --\}$. This makes the rule for P_4 simple: $(P_4, x_{2,3} = +-)$. However, merging P_2 and P_3 creates an additional problem: the possible values that $P_{2,3}$ can take depend on the value of P_1, contradicting the basic property of the rule sets, i.e. that the truth of a rule implies that a parameter is meaningful and can take any value, independently of other parameters. The solution is simple: we need to merge P_1 and $P_{2,3}$ into a bigger macro-parameter. The so-obtained parameter $P_{1,2,3}$ can take only five possible values, $\{+++, ++-, +-+, +--, -00\}$, due to restrictions imposed by internal rules between parameters that are merged.

The previous example gives an intuition of the problems involved in parameter merging. More precisely, one has to merge a (minimal) set of parameters satisfying the following consistency condition: the rule governing its meaningfulness is of the exclusive OR type, and when the rule is true, each of its values is admissible. A simple way to guarantee this constraint is to require a macro-parameter to be independent. The set of possible values of a macro-parameter is constructed by taking into account the existent relations holding between merged parameters.

Practically, once we select the first set of parameters to be merged by looking at a single, non-simple rule, we also need to merge with them all their ancestors parameters. This approach has the drawback of generating macro-parameters with large set of values. Hence, there is a trade off between space complexity of the sampling structures and time complexity of the rejection sampling approach.

6. Experimental Results and Conclusions

In this section, we present some experimental results of our method, applied on a set of 62 parameters, governed by the rules in Table 1. Details on these parameters, and on the resulting phylogenies, can be found in [1].

As we can see, some rules tend to be highly complex, so we had to merge several parameters before constructing a sampling structure. We merged parameters 25, 26, 27, 30, 31, 32, 33, 37, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49 into

$(P_1, true),\ (P_2, P_1 = +),\ (P_3, P_2 = +),\ (P_4, P_1 = +),\ (P_5, P_2 = +),\ (P_6, P_5 = +),\ (P_7, true)$
$(P_8, P_7 = +),\ (P_9, P_6 = +),\ (P_{10}, P_5 = - \vee P_6 = - \vee P_7 = +),\ (P_{11}, P_{10} = -),\ (P_{12}, P_7 = +)$
$(P_{13}, true),\ (P_{14}, P_1 = + \wedge P_8 = + \wedge P_{13} = -),\ (P_{15}, P_{12} = + \wedge P_{14} = -),\ (P_{16}, P_7 = + \wedge P_{12} = -)$
$(P_{17}, P_7 = +),\ (P_{18}, P_5 = +),\ (P_{19}, P_5 = +),\ (P_{20}, true),\ (P_{21}, true)$
$(P_{22}, (P_5 = - \vee P_6 = - \vee P_7 = +) \wedge P_{21} = -),\ (P_{23}, P_{22} = +),\ (P_{24}, P_{21} = + \vee P_{22} = +),\ (P_{25}, true)$
$(P_{26}, P_{25} = +),\ (P_{27}, P_{25} = - \vee P_{26} = +),\ (P_{28}, P_{12} = +)$
$(P_{29}, (P_{25} = - \vee P_{26} = +) \wedge (P_{25} = + \vee P_{27} = + \vee P_{28} = +)),\ (P_{30}, true),\ (P_{31}, true),\ (P_{32}, P_{31} = +)$
$(P_{33}, P_{32} = +),\ (P_{34}, true),\ (P_{35}, P_6 = + \wedge (P_{32} = + \vee P_{34} = +)),\ (P_{36}, P_{32} = + \wedge P_{14} = -)$
$(P_{37}, true),\ (P_{38}, P_{37} = +),\ (P_{39}, P_{31} = + \wedge P_{37} = +),\ (P_{40}, P_{39} = -)$
$(P_{41}, P_{31} = - \vee P_{39} = + \vee P_{40} = +),\ (P_{42}, true),\ (P_{43}, true),\ (P_{44}, P_{43} = -)$
$(P_{45}, P_{27} = + \wedge P_{44} = +),\ (P_{46}, P_{31} = + \wedge P_{44} = +),\ (P_{47}, P_{46} = +),\ (P_{48}, P_{47} = +),\ (P_{49}, P_{48} = +)$
$(P_{50}, P_{31} = - \vee P_{43} = + \vee P_{49} = +),\ (P_{51}, true),\ (P_{52}, P_{51} = +)$
$(P_{53}, (P_5 = - \vee P_6 = - \vee P_8 = +) \wedge (P_{22} = - \vee P_{51} = +)),\ (P_{54}, P_{30} = + \wedge (P_{45} = - \vee P_{46} = -))$
$(P_{55}, (P_{30} = - \vee P_{43} = + \vee P_{45} = + \vee P_{54} = -) \wedge (P_{31} = - \vee P_{32} = -) \wedge P_{42} = -)$
$(P_{56}, (P_5 = - \vee P_6 = - \vee P_8 = +) \wedge (P_{22} = - \vee P_{23} = - \vee P_{29} = -))$
$(P_{57}, (P_{30} = - \vee P_{43} = + \vee P_{45} = + \vee P_{54} = -) \wedge (P_{33} = + \vee P_{32} = - \vee P_{55} = +) \wedge P_{31} = + \wedge P_{42} = -)$
$(P_{58}, P_{46} = + \wedge (P_{47} = - \vee P_{48} = - \vee P_{49} = -) \wedge P_{57} = +),\ (P_{59}, P_{28} = -)$
$(P_{60}, P_{30} = - \vee P_{43} = + \vee P_{45} = +),\ (P_{61}, P_{45} = - \vee P_{46} = - \vee P_{47} = - \vee P_{48} = - \vee P_{13} = +)$
$(P_{62}, (P_5 = - \vee P_6 = - \vee P_{12} = -) \wedge (P_{13} = + \vee P_{14} = +))$

Table 1: Rules for the parameters in the PCM dataset [6, 1].

a large macro-parameter with 3916 possible values, and parameters 1, 2, 5, 6, 7, 8, 12, 13, 14, 21, 22 into a smaller macro-parameter with 168 values. These macro-parameters were constructed automatically, using a greedy heuristic. We started by merging parameters in the head of a selected rule and we iteratively merged more and more parameters by considering those appearing in the heads of complex rules of merged parameters. This procedure was stopped when a bound on the number of states of the macro-parameter was met. Intuitively, the problem with parameter merging is that the complex dependencies of Table 1 tend to produce a single large macro-parameter containing almost all parameters, hence we had to resort to heuristics to break this tendency.

The compact sampling structure has been constructed leaving out parameters 29, 35, 36, 53, 56, and 61 and fixing them uniformly in the set $\{+, -, 0\}$ in each attempt of the rejection sampling. On average, the rejection sampler takes slightly less that 65 trials to find an admissible language, and the average time to sample a single language is around 0.6 milliseconds. This allows one to generate 10 million languages, a reasonably large sample to extract meaningful properties of the language space, in about 2 hours. Experiments were performed on a laptop with a Core 2 Due T9300 CPU and 2 Gb of RAM.

The experimental results just presented suggest that this sampling method is feasible to deal with applications of the size of those required in [6]. Problems can arise if the number of parameters is increased and the rule structure continues to have the tendency of merging all parameters together. In such cases, we may try to improve the sampler by using a Gibbs sampling scheme [2, 7], isolating subsets of parameters that are sufficiently (but not perfectly) separated from the others.

This sampling algorithm can be used to assess the statistical significance of real world data against background noise. For instance, we can use it to check if two specific languages are significantly more similar than two languages in a pair drawn at random. It can also be used to investigate the properties of the space of admissible languages, which can shed further light on the intrinsic structure of universal grammars. In general, this method allows the grounding of PCM on firmer statistical bases.

References

[1] *Linguistics laboratory of the University of Trieste*, http://www2.units.it/linglab/.

[2] S.P. Brooks, *Markov chain Monte Carlo method and its application*, J. Roy. Statist. Soc. D **47** (1998), 69–100.

[3] N. Chomsky, *Rules and representation*, Columbia University Press, New York (1980).

[4] J. Felsenstein, *Inferring phylogenies*, Sinauer, Sunderland (2004).

[5] C. Guardiano and G. Longobardi, *Parametric comparison and language taxonomy*, in M. Batllori, M.L. Hernanz, C. Picallo and F. Roca, *Grammaticalization and Parametric Variation*, Oxford University Press, Oxford (2005), pp. 149–174.

[6] G. Longobardi and C. Guardiano, *Evidence for syntax as a signal of historical relatedness*, Lingua **119** (2009), 1679–1706.

[7] D.J.C. MacKay, *Introduction to Monte Carlo methods*, in M.I. Jordan, *Learning in graphical models*, NATO Science Series, Kluwer, U.S.A. (1998), pp. 175–204.

[8] A.M.S. McMahon and R. McMahon, *Language classification by numbers*, Oxford University Press, Oxford (2005).

Authors' addresses:

Luca Bortolussi
Dipartimento di Matematica e Informatica
Università degli Studi di Trieste
Via Valerio 12/1, 34127 Trieste, Italy
E-mail: luca@dmi.units.it

Andrea Sgarro
Dipartimento di Matematica e Informatica
Università degli Studi di Trieste
Via Valerio 12/1, 34127 Trieste, Italy
E-mail: sgarro@units.it

Received May 31, 2011
Revised August 30, 2011

Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 111–145

Splitting the Fučík Spectrum and the Number of Solutions to a Quasilinear ODE[1]

Marta García-Huidobro, Raúl Manásevich and Fabio Zanolin

Communicated by Pierpaolo Omari

Abstract. *For ϕ an increasing homeomorphism from $\mathbb{R}$ onto $\mathbb{R}$, and $f \in C(\mathbb{R})$, we consider the problem*

$$(\phi(u'))' + f(u) = 0, \quad t \in (0, L), \quad u(0) = 0 = u(L).$$

The aim is to study multiplicity of solutions by means of some generalized Pseudo Fučík spectrum (at infinity, or at zero). New insights that lead to a very precise counting of solutions are obtained by splitting these spectra into two parts, called Positive Pseudo Fučík Spectrum (PPFS) and Negative Pseudo Fučík Spectrum (NPFS) (at infinity, or at zero, respectively), in this form we can discuss separately the two cases $u'(0) > 0$ and $u'(0) < 0$.

Keywords: Fučík Spectrum, Quasilinear, p-Laplacian, Multiplicity.
MS Classification 2010: 34B15, 34A34

1. Introduction

In this paper we study the number of the solutions to the two-point (Dirichlet) boundary value problem

$$(P) \quad \begin{cases} (\phi(u'))' + f(u) = 0, & t \in (0, L), \\ u(0) = 0 = u(L), \end{cases}$$

[1]M. G.-H. was supported by Fondecyt Project Nr. 1110268.
R. M. was supported by Fondap Matemáticas Aplicadas and ISCI-Milenio grant-P05-004-F.
F. Z. was supported by Fondecyt Project Nr. 7990040. He also thanks the Facultad de Matemáticas of the Pontificia Universidad Católica de Chile for the hospitality

where ϕ is an increasing homeomorphism from $\mathbb{R}$ onto $\mathbb{R}$, and $f \in C(\mathbb{R})$.

In case that the differential operator in (P) is linear, i.e., when $\phi(s) = s$, or more generally when the differential operator is the one-dimensional p-Laplacian, i.e., when $\phi = \phi_p$ with

$$\phi_p(s) := |s|^{p-2}s, \quad s \neq 0, \quad \phi_p(0) = 0, \quad p > 1,$$

it is known that multiplicity results are obtained by assuming a suitable interaction of the nonlinearity with the Fučík spectrum of the corresponding operator. For instance, when the differential operator is linear, one is usually led to consider the limits of $\frac{f(s)}{s}$ as $s \to 0^{\pm}$ and as $s \to \pm\infty$. If these limits are all finite, say $a^{\pm}$ and $A^{\pm}$, respectively, then the number of the solutions for the boundary value problem depends on the existence of a suitable gap between the pairs (a^+, a^-) and (A^+, A^-). We recall that, in this situation, the Fučík spectrum is the set of all the pairs (μ, ν) such that the problem

$$\begin{cases} u'' + \mu u^+ - \nu u^- = 0, & t \in (0, L), \\ u(0) = 0 = u(L), \end{cases}$$

has nontrivial solutions. As it is well known this set is the union of the critical sets $\mathcal{C}_{i,j} = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 : \frac{i}{\sqrt{\mu}} + \frac{j}{\sqrt{\nu}} = \frac{L}{\pi}\}$ for i, j nonnegative integers with $|i - j| \leq 1$ (see [15]).

Here and henceforth the following notation is used: $\mathbb{R}_0^+ :=]0, +\infty[$ and $\mathbb{R}^+ := [0, +\infty[$, also $\mathbb{R}$ and $\mathbb{N}$ will denote the sets of real numbers and the set positive integers, respectively.

Similar results (see [2], [11]) have been developed for the p-Laplacian. For this case the Fučík spectrum is given by the union of the critical sets $\mathcal{C}_{i,j} = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 : \frac{i}{\mu^{1/p}} + \frac{j}{\nu^{1/p}} = \frac{L}{\pi_p}\}$ for i, j nonnegative integers with $|i - j| \leq 1$, where

$$\pi_p := 2(p-1)^{\frac{1}{p}} \int_0^1 \frac{ds}{(1 - s^p)^{\frac{1}{p}}} = 2(p-1)^{\frac{1}{p}} \frac{\pi}{p \sin(\pi/p)} ,$$

(see [10]). As for the linear differential operator, this spectrum is the set of all the pairs (μ, ν) such that the problem

$$(F_{\mu,\nu}) \qquad \begin{cases} (\phi_p(u'))' + \mu \phi_p(u^+) - \nu \phi_p(u^-) = 0, & t \in (0, L), \\ u(0) = 0 = u(L), \end{cases}$$

has nontrivial solutions, see [13]. Note that, for any $p > 1$, if $u(\cdot)$ is a nontrivial solution of the above problem, then so is $\lambda u(\cdot)$ for any positive λ. The critical sets $\mathcal{C}_{i,j}$ intersect the diagonal of the (μ, ν)-plane exactly at the sequence of the eigenvalues $\Lambda_j = (j\pi_p/L)^p$ of the differential operator $u \mapsto -(\phi_p(u'))'$ with the associated Dirichlet boundary conditions.

In the rest of this paper the following main assumptions concerning the functions ϕ and f will be considered:

(ϕ_1) *$\phi : \mathbb{R} \to \mathbb{R}$ is an increasing (not necessarily odd) bijection with $\phi(0) = 0$. We also define $\Phi(s) = \int_0^s \phi(\xi)\, dx$;*

(f_1) *$f : \mathbb{R} \to \mathbb{R}$ is continuous with $f(0) = 0$, $f(s)s > 0$ for $s > 0$ and $F(s) \to +\infty$ as $s \to \pm\infty$, where $F(s) = \int_0^s f(\xi)\, d\xi$;*

(ϕ_0) $\displaystyle\limsup_{s\to 0^\pm} \frac{\phi(\sigma s)}{\phi(s)} < +\infty$ *and* $\displaystyle\liminf_{s\to 0^\pm} \frac{\phi(\sigma s)}{\phi(s)} > 1$; *for each $\sigma > 1$,*

(ϕ_∞) $\displaystyle\limsup_{s\to \pm\infty} \frac{\phi(\sigma s)}{\phi(s)} < +\infty$ *and* $\displaystyle\liminf_{s\to \pm\infty} \frac{\phi(\sigma s)}{\phi(s)} > 1$, *for each $\sigma > 1$.*

The conditions in (ϕ_∞) were previously introduced in [16], where they were called respectively the upper and lower σ-conditions at infinity.

In connection to (P) a natural generalization of problem $(F_{\mu,\nu})$ is given by the problem

$$(P_{\mu,\nu}) \qquad \begin{cases} (\phi(u'))' + \mu\phi(u^+) - \nu\phi(u^-) = 0 & t \in (0, L), \\ u(0) = 0 = u(L). \end{cases}$$

We note that in doing this we loose homogeneity, a property which is naturally present in the definition of the Fučík spectrum for the linear or p-Laplacian cases. Nevertheless with the idea in mind that what matters is to compare some asymptotic properties of the nonlinearity f with respect to ϕ, we proposed in [18] (for the case of ϕ odd) a definition for a critical set, by using a time-mapping approach, which we called the Pseudo Fučík Spectrum (PFS) for $(P_{\mu,\nu})$ and denoted by $\mathcal{S}(\subset (\mathbb{R}_0^+)^2)$.

Next we briefly review the construction of this Spectrum. Let us consider the equation

$$(\phi(u'))' + g(u) = 0, \tag{1}$$

and recall that under suitable growth assumptions on $g \in C(\mathbb{R}, \mathbb{R})$, we can define the time-mapping

$$\mathrm{T}_g(R) := 2 \left| \int_0^R \frac{ds}{\mathcal{L}_r^{-1}(G(R) - G(s))} \right|,$$

where $G(s) = \int_0^s g(\xi)\, d\xi$ and $\mathcal{L}(s) = s\phi(s) - \int_0^s \phi(\xi)\, d\xi$. The functions $\mathcal{L}_r^{-1}$ and $\mathcal{L}_l^{-1}$, denote, respectively, the right and left inverses of $\mathcal{L}$. The number $\mathrm{T}_g(R)$

gives the distance between two consecutive zeros of a solution of (1) which attains a maximum $R > 0$ (respectively a minimum $R < 0$).

In [17], [18], for ϕ odd, we considered the problem

$$(P_\Lambda) \qquad \begin{cases} (\phi(u'))' + \Lambda\phi(u) = 0 & t \in (0, L), \\ u(0) = 0 = u(L). \end{cases}$$

Assuming that for all positive Λ the limit

$$\mathrm{T}(\Lambda) := \lim_{R\to\pm\infty} \mathrm{T}_{\Lambda\phi}(R), \tag{2}$$

exist and is a strictly decreasing functions of Λ, we considered those Λ such that $n\mathrm{T}(\Lambda) = L$ for some integer n, and called them pseudo eigenvalues for (P_Λ). With this at hand we defined the (PFS) for $(P_{\mu,\nu})$ as the set

$$\mathcal{S} = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 \mid i\mathrm{T}(\mu) + j\mathrm{T}(\nu) = L\},$$

for i, j nonnegative integers with $|i - j| \le 1$, (a somewhat more explicit description of the (PFS) will be recalled in section 2).

We also mention that we proved in [18] that for any compact set $\mathcal{K} \subset \mathbb{R}^2 \setminus \mathcal{S}$ the set of all the possible solutions for $(P_{\mu,\nu})$, with $(\mu, \nu) \in \mathcal{K}$ is a priori bounded. Furthermore, as it is easy to see, the (PFS) $\mathcal{S}$ coincides with the standard Fučík Spectrum when $\phi = \phi_p$.

We note at this point that it is more appropriate to call the set $\mathcal{S}$ a Pseudo Fučík spectrum *at infinity*. Indeed, the set $\mathcal{S}$ does not take into account any information about solutions with small norm. In Section 2 we will consider a corresponding Pseudo Fučík spectrum *at zero*. In this context we recall [20] where pseudo eigenvalues at zero were defined.

This paper is organized as follows. In Section 2 we present our main results for multiplicity of solutions for problem (P), assuming some suitable behavior at zero and at infinity of the nonlinearities involved. Indeed setting

$$\lim_{s\to 0^\pm} \frac{f(s)}{\phi(s)} = a^\pm \qquad \lim_{s\to\pm\infty} \frac{f(s)}{\phi(s)} = A^\pm,$$

our results are based on some key lemmas relating the position of the limit pairs (A^+, A^-), (a^+, a^-), in the "classical" Fučík spectrum with respect to their position in a "universal" Fučík spectrum. This comparison is possible even if the points coincide (as considered in an example at the end of this section).

In order to treat in an independent manner solutions starting with positive slope, with those starting with a negative slope, we split the Fučík spectrum (at infinity, or at zero) into two parts, that we shall call Positive Pseudo Fučík Spectrum (PPFS) and Negative Pseudo Fučík Spectrum (NPFS) (at infinity, or at zero, respectively).

In Section 3 we give a result for the strict monotonicity of time-maps, which we use to obtain the exact number of solutions for some cases, but which may also be of some independent interest. Section 4 is devoted to some examples which illustrate our results. We end the paper in Section 5 by proving some technical lemmas, of comparison type, which are needed to obtain our results.

We finish this section with an illustrative example of some of the concepts we have introduced. Let ϕ be the map defined as the odd extension to $\mathbb{R}$ of

$$s \mapsto \begin{cases} \phi_q(s), & 0 \leq s \leq 1, \\ \phi_p(s), & s \geq 1, \end{cases}$$

where $p, q > 1$ and $p \neq q$. Then the (PFS) at infinity is exactly the one corresponding to the p-laplacian, while the (PFS) at zero is the one corresponding to the q-laplacian. Both spectra look alike but for some choices of p and q it becomes clear that both spectra which quite different in scale. Thus in Figure 1, for $q = 1.3$, $p = 6.5$, and $L = \pi$, we have plotted the Fučík spectrum at zero, using different colours for the curves of type $j = i-1$, $i = j-1$ and $j = i$. Note that the square $[0, 40]^2$ of the positive quadrant contains portions of eight critical sets of the type $\mathcal{C}_{i,i}$ (the darkest ones) as well as the same number of asymmetric curves of the class $\mathcal{C}_{i,i-1}$ and of the class $\mathcal{C}_{i-1,i}$, respectively.

In Figure 2, for the same values of p, q, and L, we have plotted the Fučík spectrum at infinity. Note that now only three curves ($\mathcal{C}_{1,0}$, $\mathcal{C}_{0,1}$ and $\mathcal{C}_{1,1}$) appear in the square $[0, 40]^2$ of the positive quadrant. In Figure 3, we have put together the two previous cases in order to stress the difference in scale.

Finally let us consider the interesting situation where

$$\lim_{s \to 0^+, +\infty} \frac{f(s)}{\phi(s)} = A^+, \quad \lim_{s \to 0^-, -\infty} \frac{f(s)}{\phi(s)} = A^-,$$

i.e., a situation where the limits at 0^+ and $+\infty$ and also the limits at 0^- and $-\infty$ coincide. To make things precise, let us assume that $A^+ = 37$ and $A^- = 19$. Then the point (37,19) belongs to different non-critical regions (at zero and infinity) as it can be observed in Figures 1 and 2, respectively. Of course, this situation cannot occur for the linear or p-laplacian operator. Actually for this case there are at least 12 solutions starting with positive slope and at least 11 solutions starting with negative slope as we will see from Theorems 2.1 and 2.2 below.

2. Main results

We consider the two-point boundary value problem (P) which by convenience we recall next,

$$(P) \quad \begin{cases} (\phi(u'))' + f(u) = 0, \\ u(0) = 0 = u(L). \end{cases}$$

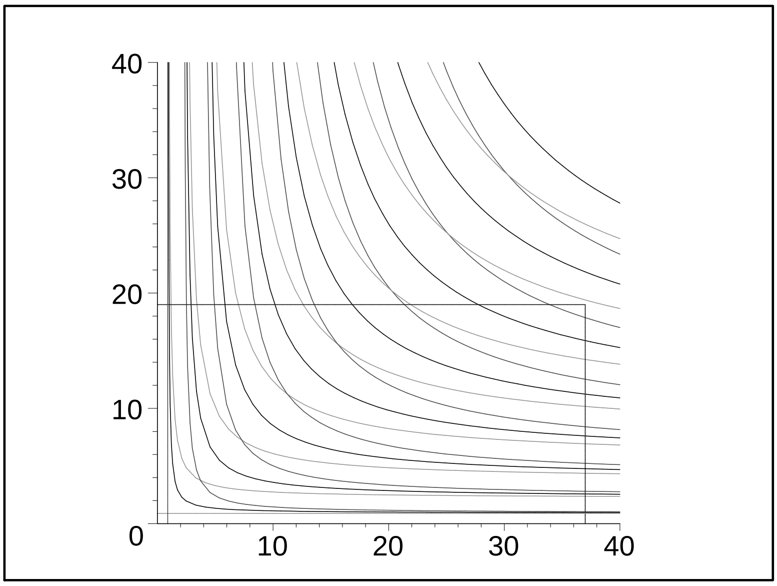


Figure 1: Fučík Spectrum for $q = 1.3$ and $L = \pi$ (PFS at zero).

We begin our analysis by assuming (ϕ_1) and (f_1) only. Under these assumptions we have that for each κ there is a unique solution $u = u(\cdot, \kappa)$ to the initial value problem

$$(\phi(u'))' + f(u) = 0, \quad u(0) = 0, \;\; u'(0) = \kappa. \tag{3}$$

This, indeed follows easily by writing equation (3) as an equivalent planar system and by using a result from [25]. Moreover, for any $\kappa \neq 0$, $u(\cdot, \kappa)$ is a nontrivial periodic solution. We will denote by $T(\kappa)$ its minimal period and by $\tau(\kappa)$ its first zero after $t = 0$.

Let us set $\Phi^*(s) = \int_0^s \phi^{-1}(\xi)\, d\xi$. Then $\mathcal{L}(s) = (\Phi^* \circ \phi)(s)$, where we recall $\mathcal{L}(s) = s\phi(s) - \Phi(s)$. As in [17], it is known that the following energy relation holds

$$\mathcal{L}(u'(t)) + F(u(t)) = \mathcal{L}(\kappa), \tag{4}$$

and it follows that $u'(\tau(\kappa)) = (\mathcal{L}_l)^{-1}(\mathcal{L}(\kappa))$ when $\kappa > 0$ and $u'(\tau(\kappa)) = (\mathcal{L}_r)^{-1}(\mathcal{L}(\kappa))$ if $\kappa < 0$, so that

$$T(\kappa) = \tau(\kappa) + \tau((\mathcal{L}_l)^{-1}(\mathcal{L}(\kappa))), \;\; \text{for } \kappa > 0$$

and

$$T(\kappa) = \tau(\kappa) + \tau((\mathcal{L}_r)^{-1}(\mathcal{L}(\kappa))), \;\; \text{for } \kappa < 0.$$

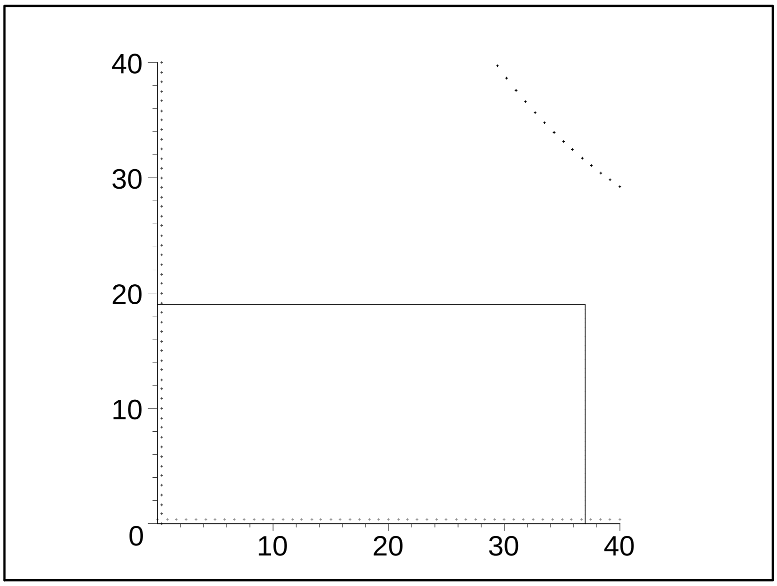


Figure 2: Fučík Spectrum for $p = 6.5$ and $L = \pi$ (PFS at infinity).

Conversely, if we fix an energy level $h > 0$, we have two solutions of (3) with $u(0) = 0$ and

$$\mathcal{L}(u'(t)) + F(u(t)) = h.$$

One of the two solutions corresponds to an initial positive slope $u'(0) = (\mathcal{L}_r)^{-1}(h)$ and the other to an initial negative slope $u'(0) = (\mathcal{L}_l)^{-1}(h)$. Clearly, any of the two solutions is a time-shift of the other. We denote by $L \cdot \mathcal{T}(h)$ (L is the length of the interval) the period of any of such solutions and also by $L \cdot x(h)$ and $L \cdot y(h)$ the distance of two consecutive zeros in an interval where the solution is respectively positive and negative.

The relationship among all the above definitions is the following:

$$x(h) + y(h) = \mathcal{T}(h) = \frac{T((\mathcal{L}_r)^{-1}(h))}{L} = \frac{T((\mathcal{L}_l)^{-1}(h))}{L},$$

and

$$x(h) = \frac{\tau((\mathcal{L}_r)^{-1}(h))}{L}, \quad y(h) = \frac{\tau((\mathcal{L}_l)^{-1}(h))}{L}.$$

By the fundamental theory of ODEs it follows that the maps $\kappa \mapsto T(\kappa)$ and $\kappa \mapsto \tau(\kappa)$, as well as $h \mapsto x(h)$ and $h \mapsto y(h)$ are continuous. Indeed, let us consider, for example, the map τ and let $\kappa_0 > 0$ be given. For any $\varepsilon > 0$, we can take $t_1, t_2 \in]0, T(\kappa_0)[$, with $\tau(\kappa_0) - \varepsilon < t_1 < \tau(\kappa_0) < t_2 < \tau(\kappa_0) + \varepsilon$, so that

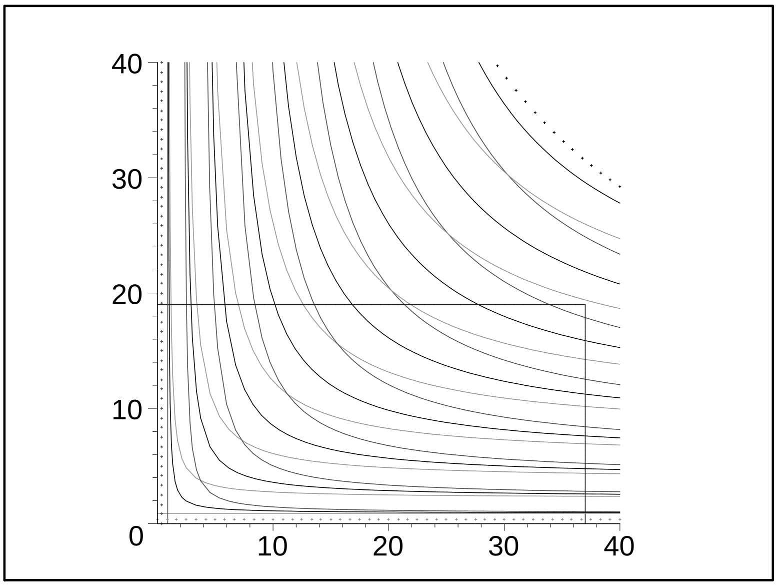


Figure 3: Overlapping the two previous Fučík Spectra.

$u(t_1, \kappa_0) > 0 > u(t_2, \kappa_0)$. Now, from the continuous dependence of the solutions on the initial data (which comes from the uniqueness of the solutions for the Cauchy problem) we have that there is $\delta > 0$ ($\delta < \kappa_0$), such that for each $\kappa \in]\kappa_0 - \delta, \kappa_0 + \delta[$ it holds that $u(t, \kappa) > 0$ for all $t \in [0, t_1]$ and $u(t_2, \kappa) < 0$. Hence, for the first zero of $u(\cdot, \kappa)$ in $]0, t_2]$, which is actually $\tau(\kappa)$, we have $\tau(\kappa) \in]t_1, t_2[\subset]\tau(\kappa_0) - \varepsilon, \tau(\kappa_0) + \varepsilon[$. The proof of the continuity for $\kappa_0 < 0$ follows the same argument and therefore we have also the continuity of $\kappa \mapsto T(\kappa)$.

In the next argument we will consider in detail the case $\kappa > 0$. Our first aim is to define a kind of "universal" Fučík spectrum in terms of the time-mappings $x(h)$ and $y(h)$. The use of universal critical sets to study boundary value problems for the equation $u'' + g(u) = 0$, was initiated in [6] and has been already developed in [4], [8] for the linear differential operator and in [9] for a Neumann problem containing the differential equation in (3). We also quote previous results which have were obtained in [3], [7], [19] and [22] for the superlinear case, and in [4], [9] for asymmetric and one-sided superlinear problems.

In this paper, as a difference with previous ones, we will "decompose" the pseudo spectrum (this will be also done for the classical Fučík spectrum) into two critical sets, corresponding respectively to solutions starting with positive and negative slope. This decomposition combined with shooting techniques will permit us to obtain precise information about the number of solutions.

Lemma 2.1. *Let ϕ and f satisfy (ϕ_1) and (f_1), respectively. Then, there exists $\kappa > 0$ such that problem (P) has a solution with $u'(0) = \kappa$ if and only if there are $n \in \mathbb{N}$ and $j \in \{0, 1\}$ such that*

$$nx(h) + (n + j - 1)y(h) = 1, \quad h = \mathcal{L}(\kappa). \tag{5}$$

Moreover, in this case, $u(\cdot)$ has exactly $2n + j - 2$ zeros in $]0, L[$ and there are n intervals in which $u > 0$ and $n + j - 1$ intervals where $u < 0$.

The proof of this lemma is straightforward and is left to the reader. Based on this result, we define the "critical lines"

$$H_i^+ = \{(x, y) \in (\mathbb{R}_0^+)^2 : \exists\, (n \in \mathbb{N}, j = 0, 1) : 2n+j-1 = i,\ nx+(n+j-1)y = 1\}.$$

Thus, H_1^+ is the half-line $x = 1$, with $y > 0$; H_2^+ is the open segment $x + y = 1$, with $x, y > 0$; H_3^+ is the open segment $2x + y = 1$, with $x, y > 0$, and so forth. The superscript "+" is to remember that these critical lines are tied up with solutions starting with positive slopes.

We will denote the set $H^+ = \cup_{i=1}^{\infty} H_i^+$ as the *Universal Positive Pseudo Fučík Spectrum* for problem $(P_{\mu,\nu})$.

Observe that if u is a solution of (3) with $u'(0) = \kappa > 0$ such that $(x(\mathcal{L}(\kappa)), y(\mathcal{L}(\kappa))) \in H_i^+$ then u is a solution of (P) having exactly $i - 1$ zeros in $]0, L[$.

Now we can divide the open first quadrant (x, y) into a countable number of open regions W_i^+ which form the complement in $(\mathbb{R}_0^+)^2$ of the critical set H^+. We label such regions as follows, see Figure 4:

- $W_1^+ = \{(x, y) : x > 1, y > 0\}$ is the part of the open first quadrant at the right hand side of H_1^+;
- $W_2^+ = \{(x, y) : 0 < x < 1, x+y > 1\}$ is the part of the open first quadrant between H_1^+ and H_2^+;
- ...
- W_i^+ is the part of the open first quadrant between H_{i-1}^+ and H_i^+, more precisely, $W_i^+ = \{(x, y) : x > 0, y > 0, kx + (k - 1)y < 1 < k(x + y)\}$ for $i = 2k$ and $W_i^+ = \{(x, y) : x > 0, y > 0, kx + ky < 1 < (k + 1)x + ky\}$ for $i = 2k + 1$.

Lemma 2.2. *Let ϕ and f satisfy (ϕ_1) and (f_1), respectively. Suppose that there are numbers $h_1, h_2 > 0$ and integers $k, \ell \geq 1$ with $k \neq \ell$ such that*

$$(x(h_1), y(h_1)) \in W_k^+, \quad (x(h_2), y(h_2)) \in W_\ell^+ .$$

Then, problem (P) has at least $|k - \ell|$ solutions with $u'(0) > 0$.

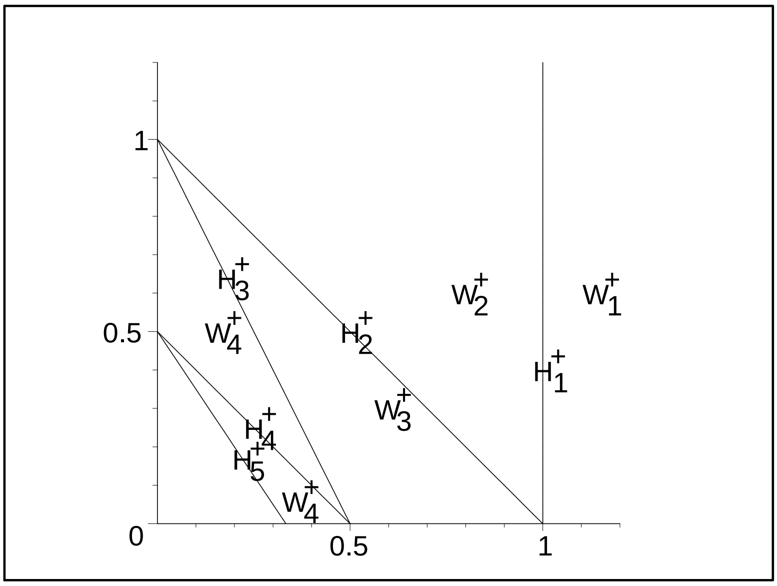


Figure 4: Universal Positive Pseudo Fučík Spectrum.

Proof. Just to fix a case, let us assume that $h_1 < h_2$ and $k < \ell$. By the assumptions, we have that the point $(x(h_1), y(h_1))$ is above H_k^+ and therefore it is above any of the H_i^+ for each $i \geq k$. On the other hand, $(x(h_2), y(h_2))$ is below $H_{\ell-1}^+$ and hence it is also below any of the H_i^+ for each $i \leq \ell - 1$.

Now, we fix an integer $i \in [k, \ell-1]$ and observe that $(x(h_1), y(h_1))$ and $(x(h_2), y(h_2))$ belong to different open components of $(\mathbb{R}_0^+)^2 \setminus H_i^+$. Since the connected set $\{(x(h), y(h))\colon\ h_1 \leq h \leq h_2\}$ has points in both components of $(\mathbb{R}_0^+)^2 \setminus H_i^+$, it must intersect H_i^+ at least once. By Lemma 2.1, this means that there is solution of (P) with positive slope at $t = 0$ and exactly $i-1$ zeros in $]0, L[$. So, all together, there are at least $\ell - k$ solutions of (P) starting with a positive slope. $\square$

Remark 2.1. *Actually, the solutions given by the lemma are such that*

$$\alpha = \min\{(\mathcal{L}_r)^{-1}(h_1), (\mathcal{L}_r)^{-1}(h_2)\} < u'(0) < \max\{(\mathcal{L}_r)^{-1}(h_1), (\mathcal{L}_r)^{-1}(h_2)\} = \beta.$$

If $\tau(\cdot)$ is strictly monotone in $]\alpha, \beta[$ and in $](\mathcal{L}_l)^{-1}((\mathcal{L}(\beta)), (\mathcal{L}_l)^{-1}((\mathcal{L}(\alpha))[$, then the number of the solutions is exactly $|k - \ell|$, for $u'(0) = \kappa$ ranging in $]\alpha, \beta[$ Indeed, if $\tau(\cdot)$ is strictly increasing (decreasing), also the maps $x(\cdot)$ and $y(\cdot)$ are strictly increasing (decreasing) with respect to h.

Our argument continues by introducing some further maps and properties that we need for the definition of the pseudo Fučík spectrum.

We deal, in first place, with the pseudo Fučík spectrum at infinity. Following [18], we assume conditions (ϕ_1) and (ϕ_∞) to hold and, moreover, that the limits

(T_∞) :
$$\mathrm{T}^{\pm}_{1,\infty}(\Lambda) = \lim_{R\to\pm\infty} \left| \int_0^R \frac{ds}{\mathcal{L}_r^{-1}(\Lambda\Phi(R) - \Lambda\Phi(s))} \right|$$

and
$$\mathrm{T}^{\pm}_{2,\infty}(\Lambda) = \lim_{R\to\pm\infty} \left| \int_0^R \frac{ds}{\mathcal{L}_l^{-1}(\Lambda\Phi(R) - \Lambda\Phi(s))} \right|$$

exist. Furthermore, we define
$$\mathrm{T}^{\pm}_{\infty}(\Lambda) = \mathrm{T}^{\pm}_{1,\infty}(\Lambda) + \mathrm{T}^{\pm}_{2,\infty}(\Lambda).$$

For each of the $\mathrm{T}^{\pm}_{\infty}$ *, we assume that either* $\mathrm{T}^{\pm}_{\infty}(\Lambda) = +\infty$ *for all* Λ*, or it is continuous and strictly decreasing with respect to* $\Lambda \in \mathbb{R}^+_0$ *. Moreover for this case, we assume that*
$$\lim_{\Lambda\to 0^+} \mathrm{T}^{\pm}_{\infty}(\Lambda) = +\infty \quad \textit{and} \quad \lim_{\Lambda\to +\infty} \mathrm{T}^{\pm}_{\infty}(\Lambda) = 0.$$

Conditions under which these hypotheses are fulfilled are given in [17], [18] and [20] for ϕ odd and T finite.

Next we define the *Positive Pseudo-Fučík Spectrum at infinity.* For $\mu, \nu > 0$ let us consider the problem ($P_{\mu,\nu}$) of the Introduction. Let us observe first that in [18], for the case of ϕ an odd function, the four numbers $\mathrm{T}^{\pm}_{i,\infty}(\Lambda)$ are all the same, finite and strictly positive, for any given Λ. Denoting this common value by $\mathrm{T}_\infty(\Lambda)$ (it corresponds to $\frac{\mathrm{T}(\Lambda)}{2}$ as defined in (2)), we can describe the (PFS) $\mathcal{S}$ (see the Introduction) as the union of the sets $\mathcal{C}_{i,i-1}$, $\mathcal{C}_{i-1,i}$, and $\mathcal{C}_{i,i}$, for $i \in \mathbb{N}$, contained in the positive (μ, ν)-quadrant, where

$$\begin{aligned} \mathcal{C}_{i,i-1} &= \{(\mu,\nu) \mid i\mathrm{T}_\infty(\mu) + (i-1)\mathrm{T}_\infty(\nu) = L/2\} \\ C_{i-1,i} &= \{(\mu,\nu) \mid (i-1)\mathrm{T}_\infty(\mu) + i\mathrm{T}_\infty(\nu) = L/2\} \\ C_{i,i} &= \{(\mu,\nu) \mid i\mathrm{T}_\infty(\mu) + i\mathrm{T}_\infty(\nu) = L/2\}. \end{aligned}$$

We want next to extend this definition to that of the Positive Pseudo-Fučík Spectrum (PPSF) at infinity, denoted by $\mathcal{S}^+(\infty)$, and where ϕ is not necessarily odd. We set
$$\mathcal{S}^+(\infty) = \cup_{i=1}^{\infty} \mathcal{C}^+_i(\infty),$$
where we consider only the case of solutions with positive (and large) slopes at $t = 0$. Here,

$$\begin{aligned} \mathcal{C}^+_1(\infty) &= \{(\mu,\nu) \in (\mathbb{R}^+_0)^2 : \ \mathrm{T}^+_\infty(\mu) = L\}, \\ \mathcal{C}^+_2(\infty) &= \{(\mu,\nu) \in (\mathbb{R}^+_0)^2 : \ \mathrm{T}^+_\infty(\mu) + \mathrm{T}^-_\infty(\nu) = L\}, \end{aligned}$$

and, in general for $i = j + k$, with $k = j - 1$ when i is odd, or $k = j$ when i is even, we have

$$\mathcal{C}_i^+(\infty) = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 : \; j\mathrm{T}_\infty^+(\mu) + k\mathrm{T}_\infty^-(\nu) = L\}.$$

We observe that the sets $\mathcal{C}_i^+(\infty)$ are all non-empty (and actually continuous curves) only in the case that both T_∞^+ and T_∞^- are finite (and hence continuous and strictly decreasing, as assumed above). In this situation the sets $\mathcal{C}_i^+(\infty)$ look similar with the corresponding $\mathcal{C}_{j,k}$ for the case ϕ odd, for $i = j + k$, with $k = j - 1$ when i is odd, or $k = j$ when i is even, see [18].

On the other hand, if T_∞^+ is finite but $\mathrm{T}_\infty^-(\Lambda) = +\infty$ for all Λ, then only $\mathcal{C}_1^+(\infty)$ is non-empty and if $\mathrm{T}_\infty^+(\Lambda) = +\infty$ for all Λ, then all the $\mathcal{C}_i^+(\infty)$'s are empty.

A qualitative description of the (PPFS) at infinity is given in the following table.

Curve	Vert. Asympt.	Horiz. Asympt.	Intersect. with Diagonal
$C_1^+(\infty)$	(Λ_1, ν)	$\emptyset$	(Λ_1, Λ_1)
$C_{2p}^+(\infty)$	(Λ_p, ν)	(μ, Λ_p)	$(\Lambda_{2p}, \Lambda_{2p})$
$C_{2p-1}^+(\infty)$	(Λ_p, ν)	(μ, Λ_{p-1})	$(\Lambda_{2p-1}, \Lambda_{2p-1})$

Arguing like in [18], we can prove the following result which corresponds to [18, Th.3.1].

Lemma 2.3. *Let f satisfy (f_1) and ϕ satisfy (ϕ_1), (ϕ_∞) and (T_∞). Then, for any compact set $\mathcal{R}$ contained in $(\mathbb{R}^+)^2 \setminus \cup_{i=1}^\infty \mathcal{C}_i^+$ there is $\kappa_{\mathcal{R}}$ such that for each $(\mu, \nu) \in \mathcal{R}$ any solution of $(P_{\mu,\nu})$ with $0 < u'(0)$ satisfies $u'(0) \le \kappa_{\mathcal{R}}$. The same result is true is we perturb the equation in $(P_{\mu,\nu})$ by adding at the right hand side a term $q(t, u)$ such that $q(t, s)/\phi(s) \to 0$ as $s \to \pm\infty$, uniformly with respect to $t \in [0, L]$.*

This result was proved in [18] for a situation corresponding to the case when T_∞^+ and T_∞^- are both finite. If T_∞^+ is finite and $\mathrm{T}_\infty^- = +\infty$, the proof is exactly the same, while if $\mathrm{T}_\infty^+ = +\infty$ we simply have a priori bounds for all the solutions starting with positive slope.

We now divide the open first quadrant (μ, ν) into a countable number of open regions $Z_i^+(\infty)$ which are the complement in $(\mathbb{R}_0^+)^2$ of the (PPFS) at infinity and label them as follows:

- $Z_1^+(\infty)$ is the part of the open first quadrant at the left hand side of $\mathcal{C}_1^+(\infty)$;
- $Z_2^+(\infty)$ is the part of the open first quadrant between $\mathcal{C}_1^+(\infty)$ and $\mathcal{C}_2^+(\infty)$;

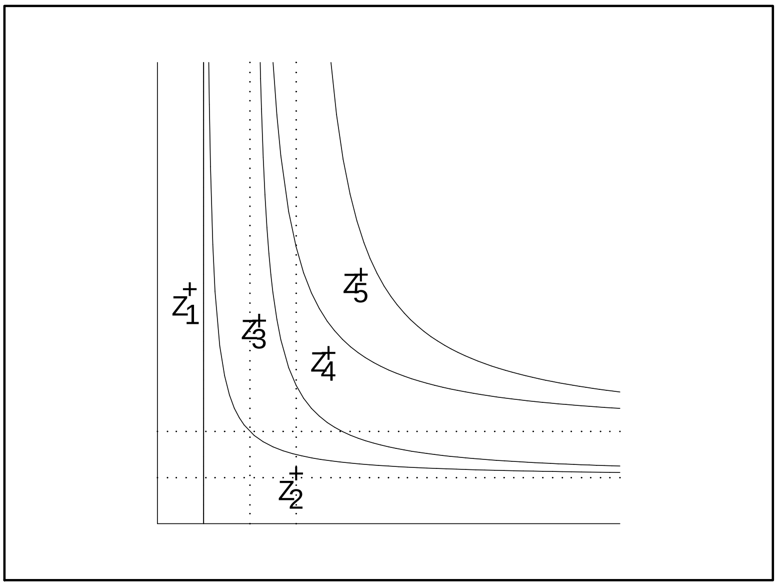


Figure 5: Positive Pseudo Fučík Spectrum at infinity.

- ...
- $Z_i^+(\infty)$ is the part of the open first quadrant between $\mathcal{C}_{i-1}^+(\infty)$ and $\mathcal{C}_i^+(\infty)$,

see Figure 5. In order to avoid a separated discussion for the cases when $\mathrm{T}_\infty^\pm$ are not finite, we point out that we have only two regions $Z_1^+(\infty)$ at the left of $\mathcal{C}_1^+(\infty)$ and $Z_2^+(\infty)$ at the right of $\mathcal{C}_1^+(\infty)$ when $\mathrm{T}_\infty^+ < +\infty$ and $\mathrm{T}_\infty^- = +\infty$. If $\mathrm{T}_\infty^+ = +\infty$ there is only one region $Z_1^+(\infty) = (\mathbb{R}_0^+)^2$.

In the special case $\phi = \phi_p$, this set is clearly a part of the standard Fučík spectrum for the one-dimensional p-Laplacian with Dirichlet boundary conditions on $]0, L[$.

Lemma 2.4. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_∞) and (T_∞). Assume that*

$$\lim_{s\to+\infty} \frac{f(s)}{\phi(s)} = A^+, \quad \lim_{s\to-\infty} \frac{f(s)}{\phi(s)} = A^-,$$

with $(A^+, A^-) \in Z_i^+(\infty)$ for some index i. Then there is $\kappa^* > 0$ such that for each $\kappa \geq \kappa^*$ it follows that $(x(\mathcal{L}(\kappa)), y(\mathcal{L}(\kappa))) \in W_i^+$.

Proof. For simplicity, we only give the proof in the case that T_∞^+ and T_∞^- are both finite. Let $\kappa > 0$ and consider the solution $u(\cdot, \kappa)$ of (3) with $u(0) = 0$ and $u'(0) = \kappa$.

First of all, we claim that the distance of two consecutive zeros of u in an interval where $u > 0$ is given by $\mathrm{T}^+_\infty(A^+) + \varepsilon_1(\kappa)$ and the distance of two consecutive zeros of u for an interval where $u < 0$ is given by $\mathrm{T}^-_\infty(A^-) + \varepsilon_2(\kappa)$, where $\varepsilon_1(\kappa) \to 0$ and $\varepsilon_2(\kappa) \to 0$ as $\kappa \to +\infty$. Hence, we have that $x(\mathcal{L}(\kappa)) - L^{-1} \cdot \mathrm{T}^+_\infty(A^+) \to 0$ and $y(\mathcal{L}(\kappa)) - L^{-1} \cdot \mathrm{T}^-_\infty(A^-) \to 0$ as $\kappa \to +\infty$.

To see this, we observe that for each $\varepsilon > 0$ there is $M_\varepsilon > 0$ such that

$$(A^+ - \varepsilon)\phi(s) \le f(s) \le (A^+ + \varepsilon)\phi(s), \quad \forall\, s \ge M_\varepsilon$$

and

$$(A^+ - \varepsilon)|\phi(s)| \le |f(s)| \le (A^+ + \varepsilon)|\phi(s)|, \quad \forall\, s \le -M_\varepsilon\,.$$

Hence, by Corollary A.1 in the appendix, and since by the energy relation (4), $\max u$, $|\min u| \to +\infty$ as $\kappa \to +\infty$, we find that

$$\mathrm{T}^+_\infty(A^+ - \varepsilon) \le \liminf_{\kappa\to+\infty} \tau(\kappa) \le \limsup_{\kappa\to+\infty} \tau(\kappa) \le \mathrm{T}^+_\infty(A^+ + \varepsilon).$$

(Here the assumption (T_∞) which guarantees the existence of the limits $\mathrm{T}^+_\infty(A^+ \pm \varepsilon)$ has been used.) Then, from these inequalities and the continuity of the function T^+_∞, we immediately obtain that

$$\lim_{\kappa\to+\infty} \tau(\kappa) = \lim_{h\to+\infty} L \cdot x(h) = \mathrm{T}^+_\infty(A^+).$$

In a completely similar manner, one can see that

$$\lim_{\kappa\to+\infty} \tau(\mathcal{L}^{-1}(\kappa)) = \lim_{h\to+\infty} L \cdot y(h) = \mathrm{T}^-_\infty(A^-),$$

concluding the proof of our claim.

Suppose next that

$$(A^+, A^-) \in Z^+_i(\infty) \quad \text{for some } \; i \in \mathbb{N}.$$

Without loss of generality we assume also that $i > 1$ and that $i = 2k$ is an even number (the case $i = 1$ is simpler as it requires only a one-sided estimate). This means that (A^+, A^-) is above the curve

$$\mathcal{C}^+_{i-1}(\infty) = \{(\mu,\nu) \in (\mathbb{R}^+_0)^2 :\; k\mathrm{T}^+_\infty(\mu) + (k-1)\mathrm{T}^-_\infty(\nu) = L\}$$

and below the curve

$$\mathcal{C}^+_i(\infty) = \{(\mu,\nu) \in (\mathbb{R}^+_0)^2 :\; k\mathrm{T}^+_\infty(\mu) + k\mathrm{T}^-_\infty(\nu) = L\}$$

and hence,

$$L < k\mathrm{T}^+_\infty(A^+) + k\mathrm{T}^-_\infty(A^-), \quad k\mathrm{T}^+_\infty(A^+) + (k-1)\mathrm{T}^-_\infty(A^-) < L.$$

Now, using the estimates in [18] for the time-mappings recalled at the beginning of this proof, we find that

$$kx(\mathcal{L}(\kappa)) + ky(\mathcal{L}(\kappa)) \to k\frac{\mathrm{T}_\infty^+(A^+)}{L} + k\frac{\mathrm{T}_\infty^-(A^-)}{L} > 1$$

and

$$kx(\mathcal{L}(\kappa)) + (k-1)y(\mathcal{L}(\kappa)) \to k\frac{\mathrm{T}_\infty^+(A^+)}{L} + (k-1)\frac{\mathrm{T}_\infty^-(A^-)}{L} < 1$$

as $\kappa \to +\infty$. This, in turns, implies that there is $\kappa^* > 0$ such that for each $\kappa > \kappa^*$ the pair $(x(\mathcal{L}(\kappa)), y(\mathcal{L}(\kappa)))$ belongs to a compact subset of W_i^+.

The case i is an odd number is clearly treated in the same way and therefore it is omitted. □

By repeating the same reasoning at zero, and using the estimates in [20], we can define *the Positive Pseudo-Fučík Spectrum at zero* $\mathcal{S}^+(0)$ as the union of a countable number of critical curves $\mathcal{C}_i^+(0)$ obtained in a similar manner as for the critical sets of the (PPFS) at infinity, but this time, with the asymptotic estimates made at zero.

More precisely let us assume the conditions (ϕ_1) and (ϕ_0), and that the limits

(T_0) :

$$\mathrm{T}_{1,0}^\pm(\Lambda) = \lim_{R\to 0^\pm}\left|\int_0^R \frac{ds}{\mathcal{L}_r^{-1}(\Lambda\Phi(R) - \Lambda\Phi(s))}\right|$$

and

$$\mathrm{T}_{2,0}^\pm(\Lambda) = \lim_{R\to 0^\pm}\left|\int_0^R \frac{ds}{\mathcal{L}_l^{-1}(\Lambda\Phi(R) - \Lambda\Phi(s))}\right|$$

exist. Furthermore, define

$$\mathrm{T}_0^\pm(\Lambda) = \mathrm{T}_{1,0}^\pm(\Lambda) + \mathrm{T}_{2,0}^\pm(\Lambda),$$

and assume that for each of the $\mathrm{T}_0^\pm$, *either* $\mathrm{T}_0^\pm(\Lambda) = +\infty$ *for all* Λ, *or it is continuous and strictly decreasing with respect to* $\Lambda \in \mathbb{R}_0^+$. *Moreover we assume that,*

$$\lim_{\Lambda\to 0^+} \mathrm{T}_0^\pm(\Lambda) = +\infty \quad \textit{and} \quad \lim_{\Lambda\to +\infty} \mathrm{T}_0^\pm(\Lambda) = 0.$$

Similarly to Lemma 2.3, we can prove the following lemma.

LEMMA 2.5. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_0) and (T_0). Then, for any compact set $\mathcal{K}$ contained in $(\mathbb{R}_0^+)^2 \setminus \cup_{i=1}^{\infty} \mathcal{C}_i^+(0)$ there is $\kappa_{\mathcal{R}}$ such that for each $(\mu,\nu) \in \mathcal{R}$ any solution of $(P_{\mu,\nu})$ with $0 < u'(0)$ satisfies $u'(0) \geq \kappa_{\mathcal{R}}$. The same result is true is we perturb the equation in $(P_{\mu,\nu})$ by adding a term $q(t,u)$ at the right hand side such that $q(t,s)/\phi(s) \to 0$ as $s \to 0$, uniformly with respect to $t \in [0,L]$.*

As before, we can define the corresponding regions in the complementary parts of the (PPFS) at zero. Accordingly, we denote by $Z_i^+(0)$ the components in the complement of the (PPFS) at zero.

We have the following result which is analogous to Lemma 2.4 and whose proof follows the same lines of that of Lemma 2.4, by using the comparison result of Lemma A.1.

LEMMA 2.6. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_0) and (T_0). Assume that*

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = a^+, \quad \lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = a^-,$$

with $(a^+, a^-) \in Z_i^+(0)$ for some index i. Then there is $\kappa_* > 0$ such that for each $0 < \kappa \leq \kappa_*$, it follows that $(x(\mathcal{L}(\kappa)), y(\mathcal{L}(\kappa))) \in W_i^+$.

We are now in a position to state our first main result.

THEOREM 2.1. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_0), (ϕ_∞) and (T_0), (T_∞). Suppose that there are positive numbers a^+, a^-, A^+, A^- such that*

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = a^+, \quad \lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = a^-$$

and

$$\lim_{s\to +\infty} \frac{f(s)}{\phi(s)} = A^+, \quad \lim_{s\to -\infty} \frac{f(s)}{\phi(s)} = A^-.$$

Assume also that there are $k, \ell \in \mathbb{N}$ with $k \neq \ell$ with $(a^+, a^-) \in Z_k^+(0)$ and $(A^+, A^-) \in Z_\ell^+(\infty)$. Then, problem (P) has at least $|k - \ell|$ solutions with $u'(0) > 0$.

Proof. The proof is a direct consequence of the Lemmas 2.4, 2.6, and 2.2, and by using the Positive Pseudo Fučík Spectrum (PPFS) at zero and at infinity. □

At this point we can establish corresponding results for solutions of (3) with $\kappa < 0$, by using a similar reasoning. To this end, with $x(h)$ and $y(h)$ as defined above, we need the following lemma which will take the place of Lemma 2.1.

LEMMA 2.7. *Let ϕ and f satisfy (ϕ_1) and (f_1), respectively. Then, there exists $\kappa < 0$ such that problem (P) has a solution with $u'(0) = \kappa$ if and only if there are $n \in \mathbb{N}$ and $j \in \{0, 1\}$ such that*

$$(n + j - 1)x(h) + ny(h) = 1, \quad h = \mathcal{L}(\kappa). \tag{6}$$

Moreover, in this case, $u(\cdot)$ has exactly $2n + j - 2$ zeros in $]0, 1[$ and also there are n intervals in which $u < 0$ and $n + j - 1$ intervals where $u > 0$.

From this result we can define, as before, the critical lines:

$$H_i^- = \{(x, y) \in (\mathbb{R}_0^+)^2 : \exists\, (n \in \mathbb{N}, j = 0, 1) : 2n+j-1 = i,\ (n+j-1)x+ny = 1\}.$$

Thus H_1^- is the half-line $y = 1$, with $x > 0$; H_2^- is the open segment $x + y = 1$, with $x, y > 0$; H_3^- is the open segment $x + 2y = 1$, with $x, y > 0$, and so forth. The superscript "-" is to remember that these critical lines are related with solutions starting with negative slopes.

We will denote the set $H^- = \cup_{i=1}^{\infty} H_i^-$ as the *Universal Negative Pseudo Fučík Spectrum*.

As before, we can divide the open first quadrant (x, y) into a countable number of open regions W_i^- which are the complement in $(\mathbb{R}_0^+)^2$ of the critical set H^-. We label such zones as follows, see Figure 6:

- $W_1^- = \{(x, y) : x > 0, y > 1\}$ is the part of the open first quadrant above H_1^- ;
- $W_2^- = \{(x, y) : x+y > 1, 0 < y < 1\}$ is the part of the open first quadrant between H_1^- and H_2^- ;
- ...
- W_i^- is the part of the open first quadrant between H_{i-1}^- and H_i^- and, more precisely, we have $W_i^- = \{(x, y) : x > 0, y > 0, (k - 1)x + ky < 1 < k(x + y)\}$ for $i = 2k$ and $W_i^- = \{(x, y) : x > 0, y > 0, kx + ky < 1 < kx + (k + 1)y\}$ for $i = 2k + 1$.

REMARK 2.2. *The set $H = H^+ \cup H^-$ has the same shape like the set $\mathcal{F}$ drawn in [4, p.874]. It represents a "universal" model of Fučík spectrum for the two-point boundary value problem in terms of time-maps. What we have done here is to distinguish the parts H^+ and H^- in H in order to treat separately the solutions with positive slope and those with negative slope. The same procedure is feasible for the standard Fučík spectrum $\mathcal{C}$ which splits as a "positive" part (concerning solutions with positive slope at $t = 0$) and a "negative" part (for the solutions with negative slope at $t = 0$).*

The following lemma is the equivalent to Lemma 2.2, it is proved similarly.

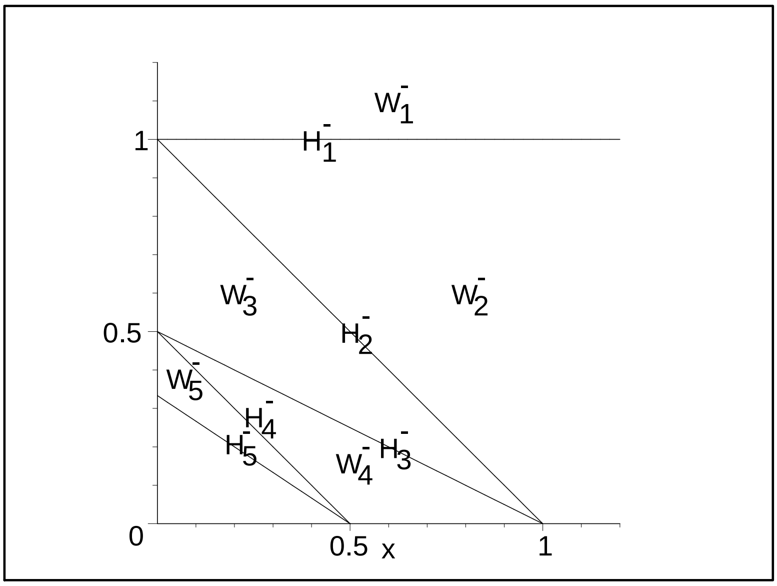


Figure 6: Universal Negative Pseudo Fučík Spectrum.

Lemma 2.8. *Let ϕ and f satisfy (ϕ_1) and (f_1), respectively. Suppose that there are numbers $h_1, h_2 > 0$ and integers $k, \ell \geq 1$ with $k \neq \ell$ such that*

$$(x(h_1), y(h_1)) \in W_k^-\,, \quad (x(h_2), y(h_2)) \in W_\ell^-\,.$$

Then, problem (P) has at least $|k - \ell|$ solutions with $u'(0) < 0$.

Remark 2.3. *In a similar way as in Remark 2.1, we find that these solutions are such that*

$$\alpha = \min\{(\mathcal{L}_l)^{-1}(h_1), (\mathcal{L}_l)^{-1}(h_2)\} < u'(0) < \max\{(\mathcal{L}_l)^{-1}(h_1), (\mathcal{L}_l)^{-1}(h_2)\} = \beta.$$

The number of solutions is exactly $|k - \ell|$, for $u'(0) = \kappa$ ranging in $]\alpha, \beta[$ if $\tau(\cdot)$ is strictly monotone in $]\alpha, \beta[$ and in $](\mathcal{L}_r)^{-1}((\mathcal{L}(\beta)), (\mathcal{L}_r)^{-1}((\mathcal{L}(\alpha))[$.

Next we define `the Negative Pseudo-Fučík Spectrum at infinity`. We set

$$\mathcal{S}^-(\infty) = \cup_{i=1}^{\infty} \mathcal{C}_i^-(\infty),$$

where we only consider solutions with negative (and large in absolute value) slopes at $t = 0$. Here,

$$\mathcal{C}_1^-(\infty) = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 : \ \mathrm{T}_\infty^-(\mu) = L\},$$
$$\mathcal{C}_2^-(\infty) = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 : \ \mathrm{T}_\infty^+(\mu) + \mathrm{T}_\infty^-(\nu) = L\},$$

and, in general, for $i = k + l$ with $k = l - 1$ (when i is odd) or $k = l$ (when i is even), we have

$$\mathcal{C}_i^-(\infty) = \{(\mu, \nu) \in (\mathbb{R}_0^+)^2 : \; k\mathrm{T}_\infty^+(\mu) + l\mathrm{T}_\infty^-(\nu) = L\}.$$

By this definition, $\mathcal{C}_i^-(\infty) = \mathcal{C}_i^+(\infty)$, when i is even.

As before the sets $\mathcal{C}_i^-(\infty)$ are all non-empty (and actually continuous curves) only in the case that both T_∞^+ and T_∞^- are finite (and hence continuous and decreasing by assumption). On the other hand, if T_∞^- is finite but $\mathrm{T}_\infty^+(\Lambda) = +\infty$ for all Λ, then only $\mathcal{C}_1^-(\infty)$ is non-empty and if $\mathrm{T}_\infty^-(\Lambda) = +\infty$ for all Λ, then all the $\mathcal{C}_i^-(\infty)$'s are empty.

A qualitative description of the (NPFS) at infinity is given in the following table.

Curve	Vert. Asympt.	Horiz. Asympt.	Intersect. with Diagonal
$C_1^-(\infty)$	(μ, Λ_1)	$\emptyset$	(Λ_1, Λ_1)
$C_{2p}^-(\infty)$	(Λ_p, ν)	(μ, Λ_p)	$(\Lambda_{2p}, \Lambda_{2p})$
$C_{2p-1}^-(\infty)$	(Λ_{p-1}, ν)	(μ, Λ_p)	$(\Lambda_{2p-1}, \Lambda_{2p-1})$

Corresponding to Lemma 2.3, we now have.

Lemma 2.9. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_∞) and (T_∞). Then, for any compact set $\mathcal{R}$ contained in $(\mathbb{R}^+)^2 \setminus \cup_{i=1}^\infty \mathcal{C}_i^-$ there is $\kappa_\mathcal{R}$ such that for each $(\mu, \nu) \in \mathcal{R}$ any solution of $(P_{\mu,\nu})$ with $u'(0) < 0$ satisfies $|u'(0)| \le \kappa_\mathcal{R}$. The same result is true is we perturb the equation in $(P_{\mu,\nu})$ by adding at the right hand side a term $q(t, u)$ such that $q(t, s)/\phi(s) \to 0$ as $s \to \pm\infty$, uniformly with respect to $t \in [0, L]$.*

We split next the open first quadrant (μ, ν) into a countable number of open regions $Z_i^-(\infty)$, which are the complement in $(\mathbb{R}_0^+)^2$ of the (NPFS) at infinity, see Figure 7, and labeled as follows:

- $Z_1^-(\infty)$ is the part of the open first quadrant below $\mathcal{C}_1^-(\infty)$;
- $Z_2^-(\infty)$ is the part of the open first quadrant between $\mathcal{C}_1^-(\infty)$ and $\mathcal{C}_2^-(\infty)$;
- ...
- $Z_i^-(\infty)$ is the part of the open first quadrant between $\mathcal{C}_{i-1}^-(\infty)$ and $\mathcal{C}_i^-(\infty)$.

Also as before and in order to avoid to discuss separately the cases when the $\mathrm{T}_\infty^\pm$ are not finite, we point out that there are only two regions $Z_1^-(\infty)$ below $\mathcal{C}_1^-(\infty)$ and $Z_2^-(\infty)$ above $\mathcal{C}_1^-(\infty)$ when $\mathrm{T}_\infty^- < +\infty$ and $\mathrm{T}_\infty^+ = +\infty$. If $\mathrm{T}_\infty^- = +\infty$ there is only one region $Z_1^-(\infty) = (\mathbb{R}_0^+)^2$.

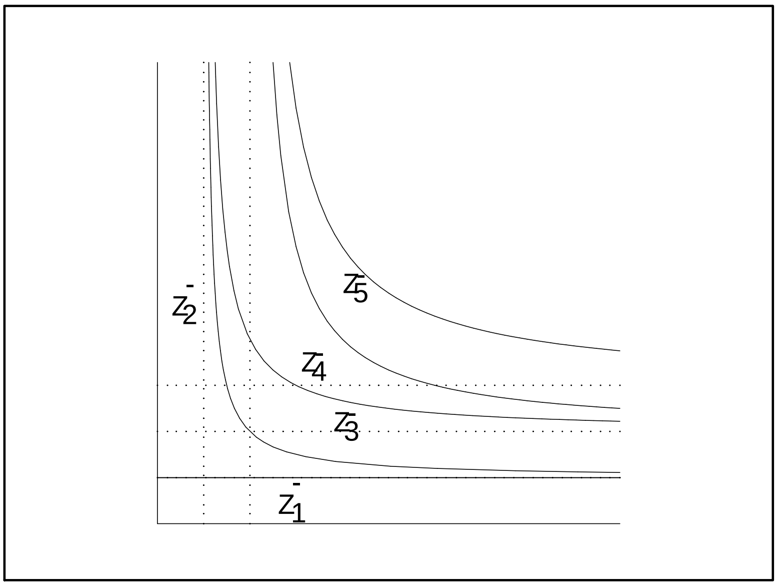


Figure 7: Negative Pseudo Fučík Spectrum at infinity.

The following lemma is the counterpart of Lemma 2.4 and it is proved in the same form.

LEMMA 2.10. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_∞) and (T_∞). Assume that*

$$\lim_{s\to+\infty}\frac{f(s)}{\phi(s)}=B^+,\quad \lim_{s\to-\infty}\frac{f(s)}{\phi(s)}=B^-,$$

with $(B^+,B^-)\in Z_i^-(\infty)$ for some index i. Then there is $\kappa^*>0$ such that for each $\kappa\le-\kappa^*$ it follows that $(x(\mathcal{L}(\kappa)),y(\mathcal{L}(\kappa)))\in W_i^-$.

REMARK 2.4. *The critical set $\mathcal{S}$ defined in [18], that is the (PFS) at infinity, is exactly $\mathcal{S}^+(\infty)\cup\mathcal{S}^-(\infty)$.*

By repeating the argument, and using estimates like in [5] we can define `the Negative Pseudo-Fučík Spectrum at zero` $\mathcal{C}^-(0)$ as the union of a countable number of critical curves $\mathcal{C}_i^-(0)$ defined in a similar manner as the critical sets of the (NPFS) at infinity, but, this time, with the estimates made at zero. Also we can define the corresponding zones $Z_i^-(0)$ as the complementary parts of the (NPFS) at zero.

Similarly to Lemmas 2.9 and 2.10, we now have.

LEMMA 2.11. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_0) and (T_0). Then, for any compact set $\mathcal{K}$ contained in $(\mathbb{R}^+)^2\setminus\cup_{i=1}^\infty\mathcal{C}_i^-(0)$ there is $\kappa_{\mathcal{R}}$ such that for* *each* $(\mu,\nu) \in \mathcal{R}$ *any solution of* $(P_{\mu,\nu})$ *with* $u'(0) < 0$ *satisfies* $|u'(0)| \geq \kappa_{\mathcal{R}}$. *The same result is true is we perturb the equation in* $(P_{\mu,\nu})$ *by adding at the right hand side a term* $q(t,u)$ *such that* $q(t,s)/\phi(s) \to 0$ *as* $s \to 0$, *uniformly with respect to* $t \in [0,L]$.

LEMMA 2.12. *Let* f *satisfy* (f_1) *and let* ϕ *satisfy* (ϕ_1), (ϕ_0) *and* (T_0). *Assume that*

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = b^+, \quad \lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = b^-,$$

with $(b^+,b^-) \in Z_i^-(0)$ *for some index* i. *Then there is* $\kappa_* > 0$ *such that for each* $-\kappa_* \leq \kappa < 0$ *it follows that* $(x(\mathcal{L}(\kappa)), y(\mathcal{L}(\kappa))) \in W_i^-$.

We thus have reached a point where we can establish and prove our second main theorem.

THEOREM 2.2. *Let* f *satisfy* (f_1) *and let* ϕ *satisfy* (ϕ_1), (ϕ_0), (ϕ_∞) *and* (T_0), (T_∞). *Suppose that there are positive numbers* b^+, b^-, B^+, B^- *such that*

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = b^+, \quad \lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = b^-$$

and

$$\lim_{s\to +\infty} \frac{f(s)}{\phi(s)} = B^+, \quad \lim_{s\to -\infty} \frac{f(s)}{\phi(s)} = B^-.$$

Assume also that there are $k,\ell \in \mathbb{N}$ *with* $k \neq \ell$ *with* $(b^+,b^-) \in Z_k^-(0)$ *and* $(B^+,B^-) \in Z_\ell^-(\infty)$. *Then, problem* (P) *has at least* $|k-\ell|$ *solutions with* $u'(0) < 0$.

Proof. Direct consequence of Lemmas 2.11, 2.12, and 2.8. □

Note that given a pair $(\mu,\nu) \in (\mathbb{R}_0^+)^2$, not belonging to the critical set at zero $\mathcal{S}^+(0)$, we can determine the region $Z_i^+(0)$ to which it belongs, by the following criterion. Let us set

$$\rho(0) = \frac{L}{T_0^+(\mu) + T_0^-(\nu)}$$

and observe that there is only one integer, say j, belonging to the open interval

$$\left]\rho(0) - \frac{T_0^+(\mu)}{T_0^+(\mu)+T_0^-(\nu)}, \rho(0) + \frac{T_0^-(\nu)}{T_0^+(\mu)+T_0^-(\nu)}\right[.$$

Then, we have that (μ,ν) belongs to $Z_{2j}^+(0)$ or $Z_{2j+1}^+(0)$, according to whether $j > \rho(0)$ or $j < \rho(0)$.

Avoiding to consider the semi-trivial cases in which one or both of the two maps $T_0^{\pm}$ are infinite, we observe that $\rho(0)$ as a function of μ or ν is continuous, strictly increasing and such that $\rho(0) \to 0$ as $\mu \to 0^+$ or $\nu \to 0^+$. Moreover $\rho(0) \to \frac{L}{T_0^-(\nu)}$ as $\mu \to +\infty$ and $\rho(0) \to \frac{L}{T_0^+(\mu)}$ as $\mu \to +\infty$. On the other hand, for $\alpha(0) = \frac{T_0^+(\mu)}{T_0^+(\mu)+T_0^-(\nu)}$, we have that $0 < \alpha(0) < 1$, with $\alpha(0)$ decreasing in μ and increasing in ν and such that $\alpha(0) \to 1$ as $\mu \to 0^+$, or $\mu \to +\infty$ and $\alpha(0) \to 0$ as $\mu \to +\infty$, or $\nu \to 0^+$.

Clearly, the same procedure can be followed in order to determine the region $Z_i^+(\infty)$ to which the pair (μ, ν) belongs.

Similarly, given a pair $(\mu, \nu) \in (\mathbb{R}_0^+)^2$ not belonging to the critical set at zero $\mathcal{S}^-(0)$, we can determine the region $Z_i^-(0)$ it belongs, by the following criterion. Let us set

$$\rho(0) = \frac{L}{T_0^+(\mu) + T_0^-(\nu)}\,,$$

and as before observe that there is only one integer, say j, belonging to the open interval

$$\left]\rho(0) - \frac{T_0^-(\nu)}{T_0^+(\mu) + T_0^-(\nu)}\,, \rho(0) + \frac{T_0^+(\mu)}{T_0^+(\mu) + T_0^-(\nu)}\right[\,.$$

Then, we have that (μ, ν) belongs to $Z_{2j}^-(0)$ or $Z_{2j+1}^-(0)$, according to whether $j > \rho^+(0)$ or $j < \rho^+(0)$.

The same procedure can be followed in order to determine the region $Z_i^-(\infty)$ the pair (μ, ν) belongs to.

Remark 2.5. *In the proof of the main theorems, it is not important that the limiting pairs (a^+, a^-) and (A^+, A^-) in Theorem 2.1, or (b^+, b^-), and (B^+, B^-) in Theorem 2.2 be in "nonresonance" zones out of the (PFS) sets. What really matters is that there is a suitable gap between the corresponding positions of these limits at zero and at infinity with respect to the universal Fučík spectrum. This situation reminds the occurrence of a twist condition between zero and infinity which permits the application of the Poincaré-Birkhoff fixed point theorem in the periodic case [11].*

In view of the above remark, the following rule can be given.

The set $\mathcal{U}^+ := \{H_i^+, W_i^+ : i \in \mathbb{N}\}$ represents a partition of $(\mathbb{R}_0^+)^2$. Given two points $P, Q \in (\mathbb{R}_0^+)^2$, we have that the number of transverse intersections between the *open* segment $]P, Q[$ and the critical lines $\cup_{i=1}^{\infty} H_i^+$ depends only on which of the classes of the above partition P and Q belong to (notice that, in this way, if P and Q are both on the same critical line H_i^+, then the intersection counts like zero).

The sets $\mathcal{M}^+(0) := \{\mathcal{C}_i^+(0), Z_i^+(0) : i \in \mathbb{N}\}$ and $\mathcal{M}^+(\infty) := \{\mathcal{C}_i^+(\infty), Z_i^+(\infty) : i \in \mathbb{N}\}$ also determine a partition of $(\mathbb{R}_0^+)^2$. We can define now the maps

$$\mathcal{M}^+(0) \to \mathcal{U}^+, \quad \mathcal{M}^+(\infty) \to \mathcal{U}^+$$

by

$$\mathcal{C}_i^+(0),\ \mathcal{C}_i^+(\infty) \mapsto H_i^+ \quad \text{and} \quad Z_i^+(0),\ Z_i^+(\infty) \mapsto W_i^+ .$$

Note that these maps make a correspondence between curves to lines and open sets to open sets.

Take a point $(a,b) \in (\cup_{i=1}^\infty \mathcal{C}_i^+(0)) \cup (\cup_{i=1}^\infty Z_i^+(0))$. We can associate to (a,b) the set in $\mathcal{M}^+(0)$ to which it belongs and hence, to this one, the set in $\mathcal{U}^+$ which is associated to it via the above map. Call this set $[(a,b)]$. Similarly, given a point $(A,B) \in (\cup_{i=1}^\infty \mathcal{C}_i^+(\infty)) \cup (\cup_{i=1}^\infty Z_i^+(\infty))$ we can map it to a set $[(A,B)] \in \mathcal{U}^+$, where $[(A,B)]$ is the set corresponding to that one in $\mathcal{M}^+(\infty)$ to which (A,B) belongs. Since, as observed before, the number of transverse intersections of the open segment $]P,Q[$ with the set $\cup_{i=1}^\infty H_i^+$ is the same for each $P \in [(a,b)]$ and $Q \in [(A,B)]$, we have that this number is well determined by the initial choice of the pairs (a,b) and (A,B). We denote this number by $i^+[(a,b),(A,B)]$ and call it the *positive intersection index* for the pairs (a,b) and (A,B).

We remark that this definition requires that the pairs (a,b) and (A,B) are "thought" in relation with the (PPFS) at zero and at infinity, respectively.

In a similar manner, one can define the partitions $\mathcal{U}^-$, $\mathcal{M}^-(0)$, $\mathcal{M}^-(\infty)$ of $(\mathbb{R}_0^+)^2$ and the maps

$$\mathcal{C}_i^-(0),\ \mathcal{C}_i^-(\infty) \mapsto H_i^- \quad \text{and} \quad Z_i^-(0),\ Z_i^-(\infty) \mapsto W_i^- ,$$

in order to define an index $i^-[(c,d),(C,D)]$ as the *negative intersection index* for the pairs (c,d) and (C,D), where (c,d) is related to the (NPFS) at zero and (CD) to the (NPFS) at infinity.

We point out that this procedure works also in the "degenerate" case in which some of the $\mathrm{T}_0^\pm(\Lambda)$ or $\mathrm{T}_\infty^\pm(\Lambda)$ considered in (T_0) and in (T_∞) is constantly equal to infinity. In this situation, some of the maps $\mathcal{M}^\pm(0), \mathcal{M}^\pm(0) \to \mathcal{U}^\pm$ will be not surjective, but the definition of the intersection indexes is well posed too.

Then, we have:

Theorem 2.3. *Let f satisfy (f_1) and let ϕ satisfy (ϕ_1), (ϕ_0), (ϕ_∞) and (T_0), (T_∞). Suppose that there are positive numbers d^+, d^-, D^+, D^- such that*

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = d^+, \quad \lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = d^-$$

and

$$\lim_{s\to +\infty} \frac{f(s)}{\phi(s)} = D^+, \quad \lim_{s\to -\infty} \frac{f(s)}{\phi(s)} = D^-.$$

Then, problem (P) *has at least so many solutions with* $u'(0) > 0$ *like the positive intersection index* $i^+[(d^+, d^-), (D^+, D^-)]$ *and at least so many solutions with* $u'(0) < 0$ *like the negative intersection index* $i^-[(d^+, d^-), (D^+, D^-)]$.

3. A result for strictly monotone time-mappings

We present in this section some suitable conditions under which we have a strictly monotone time-map. Hence, according to the remarks after Lemmas 2.2 and 2.8, they can be applied to obtain an exact number of solutions, .

We shall confine ourselves only to the consideration of $\tau(\kappa)$ for $\kappa > 0$. The situation in which $\kappa < 0$ is completely symmetric and therefore can be discussed using the same arguments.

For simplicity, we suppose that $\phi : \mathbb{R} \to \mathbb{R}$ is an odd increasing homeomorphism which is of class C^1 in $\mathbb{R}_0^+$ and such that

$$(\phi_2) \qquad \lim_{s \to 0^+} s^2 \phi'(s) = 0$$

and we also assume that $f : \mathbb{R} \to \mathbb{R}$ satisfy (f_1). Let $u(\cdot)$ be a solution of (3) for some $\kappa > 0$.

From the energy relation (4) we can compute the distance $\tau(\kappa)$ of two consecutive zeros of u in an interval where $u > 0$ by the formula

$$\tau(\kappa) = 2 \int_0^R \frac{ds}{\mathcal{L}_r^{-1}(F(R) - F(s))}\,, \quad \text{with } \ F(R) = \mathcal{L}(\kappa),\ R > 0,$$

where the number "2" comes from the fact that, by the symmetry of ϕ, we have that the time from $t = 0$ to the point of maximum of u is the same like the time from the point of maximum of u and $t = \tau(\kappa)$.

For convenience in the next proof, we also introduce the function

$$\Psi(s) := \phi'(\mathcal{L}_r^{-1}(s))(\mathcal{L}_r^{-1}(s))^2,$$

which, by the assumptions on ϕ, is defined and continuous for every $s > 0$ and, by (ϕ_2) can be extended by continuity to the origin, putting $\Psi(0) = 0$.

Then, the following result can be proved:

Lemma 3.1. *Assume* (ϕ_2) *and suppose that the function* Ψ *is convex (respectively, concave) in* $\mathbb{R}^+$. *Then, the map* $\kappa \mapsto \tau(\kappa)$ *is strictly decreasing (respectively, strictly increasing) if the map* $s \mapsto sf(s) - \Psi(F(s))$ *is strictly increasing (respectively, strictly decreasing).*

Proof. First of all, via the change of variables $s = Rt$, we write the integral for $\tau(\kappa)$ as

$$\tau(\kappa) = 2 \int_0^1 \frac{R\,dt}{\mathcal{L}_r^{-1}(F(R) - F(Rt))}\,,$$

so that, in order to prove that $\tau(\cdot)$ is strictly decreasing (respectively, strictly increasing) for κ in a certain interval, say $\kappa \in]\alpha, \beta[$, with $0 \le \alpha < \beta \le +\infty$, we may prove that, for each $t \in]0,1[$ the map $R \mapsto (\mathcal{L}_r^{-1}(F(R) - F(Rt)))/R$ is strictly increasing (or strictly decreasing) on $]\alpha_1, \beta_1[$, where we have set $\alpha_1 = F_r^{-1}(\mathcal{L}(\alpha))$ and $\beta_1 = F_r^{-1}(\mathcal{L}(\beta))$. To do this, we compute the derivative with respect to R for any fixed $t \in]0,1[$. Then, at the numerator we find

$$\frac{Rf(R) - Rtf(Rt)}{\mathcal{L}'(\mathcal{L}_r^{-1}(F(R) - F(Rt)))} - \mathcal{L}_r^{-1}(F(R) - F(Rt)).$$

Recalling that $\mathcal{L}'(\xi) = \xi\phi'(\xi)$ and the definition of Ψ, we conclude that τ is strictly decreasing (strictly increasing) in $]\alpha, \beta[$, if

$$Rf(R) - Rtf(Rt) - \Psi(F(R) - F(Rt)) > 0 \tag{7}$$

holds every $R \in]\alpha_1, \beta_1[$ and all $t \in]0,1[$. Now, according to the assumption, we have that Ψ is a convex function and recall also that $\Psi(0) = 0$. From this, it follows that $\Psi(a+b) \ge \Psi(a) + \Psi(b)$ for all $a, b \ge 0$ and therefore, $\Psi(x-y) \le \Psi(x) - \Psi(y)$ for all $0 < y < x$. Hence, we have that $\Psi(F(R) - F(Rt)) \le \Psi(F(R)) - \Psi(F(Rt))$ holds for every $R \in]\alpha_1, \beta_1[$ and all $t \in]0,1[$ and therefore, in order to prove (7) it will be sufficient to prove that

$$Rf(R) - Rtf(Rt) - [\Psi(F(R)) - \Psi(F(Rt))] > 0$$

holds for every $R \in]\alpha_1, \beta_1[$ and all $t \in]0,1[$.

Now, to have this last inequality satisfied it will be enough to have that the function $R \mapsto Rf(R) - \Psi(F(R))$ is strictly increasing on the interval $]0, \beta_1[$.

In case that the function Ψ is concave, we use the inequality $\Psi(x-y) \ge \Psi(x) - \Psi(y)$ for all $0 < y < x$ and obtain that the time-mapping is strictly increasing on $]\alpha, \beta[$ provided that the map $R \mapsto Rf(R) - \Psi(F(R))$ is strictly decreasing in the interval $]0, \beta_1[$. □

Note that if ϕ is of class C^2 in $\mathbb{R}_0^+$, then the map Ψ is convex (respectively, concave) provided that $s\phi''(s)/\phi'(s)$ is increasing (respectively, decreasing) on $\mathbb{R}_0^+$. In the special case of $\phi = \phi_p$, for some $p > 1$, we have that $s\phi''(s)/\phi'(s)$ is a positive constant.

An elementary application of Lemma 3.1 is the following:

Corollary 3.1. *For $\phi = \phi_p$, with $p > 1$, the following holds: The map $\kappa \mapsto \tau(\kappa)$ is strictly decreasing (respectively, strictly increasing) if the map $s \mapsto \frac{f(s)}{s^{p-1}}$ is strictly increasing (respectively, strictly decreasing).*

Proof. In this case, by a direct computation, we have that $\Psi(s) = ps$, so that all the assumptions on Ψ are satisfied and the auxiliary function $s \mapsto sf(s) - \Psi(F(s))$ takes the form of $sf(s) - pF(s)$, which, in turn, is strictly increasing (respectively, strictly decreasing) if the map $s \mapsto \frac{f(s)}{s^{p-1}}$ is strictly increasing (respectively, strictly decreasing), too. □

Corollary 3.1 extends a classical result obtained by Opial [24] for the case $p = 2$. Other estimates for the time-mappings associated to the ϕ_p or the ϕ-operators can be found in [17], [18], [20] and [23].

A variant of Lemma 3.1 which gives as a consequence Corollary 3.1 as well, is the following.

LEMMA 3.2. *Assume* (ϕ_2) *and suppose that there is a constant* $\theta > 0$ *such that* $\Psi(s) \leq \theta s$ *(respectively* $\Psi(s) \geq \theta s$*) for all* $s \geq 0$*. Then, the map* $\kappa \mapsto \tau(\kappa)$ *is strictly decreasing (respectively, strictly increasing) if the map* $s \mapsto \frac{f(s)}{s^{\theta-1}}$ *is strictly increasing (respectively, strictly decreasing).*

This result still admits a little variant, in the sense that if we know that $\Psi(s) < \theta s$ for $s > 0$, then it will be sufficient to assume $s \mapsto \frac{f(s)}{s^{\theta-1}}$ increasing (weakly) in order to have τ strictly decreasing in κ and, conversely, if $\Psi(s) > \theta s$ for $s > 0$, then τ is strictly increasing when $s \mapsto \frac{f(s)}{s^{\theta-1}}$ is decreasing (weakly).

As a final remark for this section, we observe that all the results presented here can be extended, modulo suitable changes, if we assume that there are $-\infty \leq a < 0 < b \leq +\infty$ and $-\infty \leq c < 0 < d \leq +\infty$, such that $\phi :]a, b[\to]c, d[$ is a strictly increasing bijection with $\phi(0) = 0$ and also we suppose that ϕ is of class C^1 in $\mathbb{R}_0^+$ with $s^2\phi'(s) \to 0$, as $s \to 0^{\pm}$.

In this situation, however, the results of monotonicity for the time-mapping will have their range of validity only for a suitable neighborhood of the origin. More precisely, for $u'(0) = \kappa > 0$, we have to take $\kappa \in]0, \beta[$, with $\beta = \min\{b, \mathcal{L}_r^{-1}(\mathcal{L}(a))\}$ and the corresponding R's will vary in $]R_l^{-1}(\mathcal{L}(\beta)), F_r^{-1}(\mathcal{L}(\beta)[$. For $u'(0) = \kappa < 0$, we have to take $\kappa \in]\alpha, 0[$, with $\alpha = \max\{a, \mathcal{L}_l^{-1}(\mathcal{L}(b))\}$ and the corresponding R's will vary in $]F_l^{-1}(\mathcal{L}(\alpha), F_r^{-1}(\mathcal{L}(\alpha))[$.

4. Examples

In this section we illustrate some of our results through simple examples.

Example 1. Let $p,\ q > 1$ and define the homeomorphism ϕ by

$$\phi(s) = \begin{cases} \phi_p(s), & \text{for } s \geq 0, \\ \phi_q(s), & \text{for } s \leq 0. \end{cases} \tag{8}$$

Also, let us denote by B the well known beta function (see for example [1, p. 258])

$$B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1} dx,$$

which we recall is convergent for $m,\ n>0$. Then, for any positive h, we find that

$$L\cdot x(h)=\Big(\frac{p}{\mu}\Big)^{\frac{1}{p}}\left[\Big(\frac{1}{p'}\Big)^{\frac{1}{p}}\int_0^1\frac{dz}{(1-z^p)^{\frac{1}{p}}}+\Big(\frac{1}{q'}\Big)^{\frac{1}{q}}\int_0^1\frac{dz}{(1-z^p)^{\frac{1}{q}}}(\mu h)^{\frac{1}{p}-\frac{1}{q}}\right]$$
$$=\frac{\pi_p}{2\mu^{\frac{1}{p}}}+\frac{B(\frac{1}{p},\frac{1}{q'})}{p^{\frac{1}{p'}}(q')^{\frac{1}{q}}}\cdot\frac{h^{\frac{1}{p}-\frac{1}{q}}}{\mu^{\frac{1}{q}}}$$

and

$$L\cdot y(h)=\Big(\frac{q}{\nu}\Big)^{\frac{1}{q}}\left[\Big(\frac{1}{q'}\Big)^{\frac{1}{q}}\int_0^1\frac{dz}{(1-z^q)^{\frac{1}{q}}}+\Big(\frac{1}{p'}\Big)^{\frac{1}{p}}\int_0^1\frac{dz}{(1-z^q)^{\frac{1}{p}}}(\nu h)^{\frac{1}{q}-\frac{1}{p}}\right]$$
$$=\frac{\pi_q}{2\nu^{\frac{1}{q}}}+\frac{B(\frac{1}{q},\frac{1}{p'})}{q^{\frac{1}{q'}}(p')^{\frac{1}{p}}}\cdot\frac{h^{\frac{1}{q}-\frac{1}{p}}}{\nu^{\frac{1}{p}}},$$

where $x(h)$ and $y(h)$ are referred to the equation in $(P_{\mu,\nu})$. We also have

$$\mathrm{T}^+_{1,\infty}(\Lambda)=\Big(\frac{p}{\Lambda}\Big)^{\frac{1}{p}}\left[\Big(\frac{1}{p'}\Big)^{\frac{1}{p}}\int_0^1\frac{dz}{(1-z^p)^{\frac{1}{p}}}\right]=\frac{\pi_p}{2\Lambda^{\frac{1}{p}}},$$
$$\mathrm{T}^+_{2,\infty}(\Lambda)=\begin{cases}0, & \text{if } q<p,\\ \dfrac{\pi_p}{2\Lambda^{\frac{1}{p}}}, & \text{if } q=p,\\ +\infty, & \text{if } q>p,\end{cases}$$
$$\mathrm{T}^-_{2,\infty}(\Lambda)=\Big(\frac{q}{\Lambda}\Big)^{\frac{1}{q}}\left[\Big(\frac{1}{q'}\Big)^{\frac{1}{q}}\int_0^1\frac{dz}{(1-z^q)^{\frac{1}{q}}}\right]=\frac{\pi_q}{2\Lambda^{\frac{1}{q}}},$$
$$\mathrm{T}^-_{1,\infty}(\Lambda)=\begin{cases}+\infty, & \text{if } q<p,\\ \dfrac{\pi_q}{2\Lambda^{\frac{1}{q}}}, & \text{if } q=p,\\ 0, & \text{if } q>p.\end{cases}$$

Finally,

$$\mathrm{T}^+_{\infty}(\Lambda)=\begin{cases}\dfrac{\pi_p}{2\Lambda^{\frac{1}{p}}}, & \text{if } q<p,\\[2ex] \dfrac{\pi_p}{\Lambda^{\frac{1}{p}}}, & \text{if } q=p,\\[2ex] +\infty, & \text{if } q>p\end{cases}\quad\text{and}\quad \mathrm{T}^-_{\infty}(\Lambda)=\begin{cases}+\infty, & \text{if } q<p,\\[2ex] \dfrac{\pi_p}{\Lambda^{\frac{1}{p}}}, & \text{if } q=p,\\[2ex] \dfrac{\pi_p}{2\Lambda^{\frac{1}{p}}}, & \text{if } q>p.\end{cases}$$

We note that when $p=q$, the (PPFS) at infinity is $\mathcal{S}^+(\infty)=\cup_{i=1}^{\infty}\mathcal{C}^+_i(\infty)$, where

$$\mathcal{C}^+_{2j-1}(\infty)=\left\{(\mu,\nu)\in(\mathbb{R}^+_0)^2:\ \frac{j}{\mu^{\frac{1}{p}}}+\frac{j-1}{\nu^{\frac{1}{p}}}=\frac{L}{\pi_p}\right\}$$

and

$$\mathcal{C}_{2j}^{+}(\infty) = \left\{ (\mu,\nu) \in (\mathbb{R}_0^+)^2 : \frac{j}{\mu^{\frac{1}{p}}} + \frac{j}{\nu^{\frac{1}{p}}} = \frac{L}{\pi_p} \right\}$$

This set is clearly a part of the standard Fučík spectrum for the one-dimensional p-Laplacian with respect to the Dirichlet boundary conditions on $]0, L[$. On the other hand, when $p > q$, the (PPFS) at infinity is actually made only by the line

$$\mathcal{C}_1^{+}(\infty) = \left\{ (\mu,\nu) \in (\mathbb{R}_0^+)^2 : \mu = \left(\frac{\pi_p}{2L}\right)^p \right\}$$

and when $p < q$, it is empty.

Similarly, we have that

$$\mathrm{T}_{1,0}^{+}(\Lambda) = \frac{\pi_p}{2\Lambda^{\frac{1}{p}}}, \quad \mathrm{T}_{2,0}^{+}(\Lambda) = \begin{cases} \infty\,, & \text{if } q < p, \\ \dfrac{\pi_p}{2\Lambda^{\frac{1}{p}}}\,, & \text{if } q = p, \\ 0\,, & \text{if } q > p, \end{cases}$$

$$\mathrm{T}_{2,0}^{-}(\Lambda) = \frac{\pi_q}{2\Lambda^{\frac{1}{q}}}, \quad \mathrm{T}_{1,0}^{-}(\Lambda) = \begin{cases} 0\,, & \text{if } q < p, \\ \dfrac{\pi_q}{2\Lambda^{\frac{1}{q}}}\,, & \text{if } q = p, \\ +\infty\,, & \text{if } q > p. \end{cases}$$

Hence,

$$\mathrm{T}_0^{+}(\Lambda) = \begin{cases} +\infty\,, & \text{if } q < p, \\ \dfrac{\pi_p}{\Lambda^{\frac{1}{p}}}\,, & \text{if } q = p, \\ \dfrac{\pi_p}{2\Lambda^{\frac{1}{p}}}\,, & \text{if } q > p \end{cases} \quad \text{and} \quad \mathrm{T}_0^{-}(\Lambda) = \begin{cases} \dfrac{\pi_q}{2\Lambda^{\frac{1}{q}}}\,, & \text{if } q < p, \\ \dfrac{\pi_p}{\Lambda^{\frac{1}{p}}}\,, & \text{if } q = p, \\ +\infty\,, & \text{if } q > p. \end{cases}$$

Clearly, when $p = q$, the (PPFS) at zero is the same as the one at infinity, while, when $p > q$ it is empty and, when $p < q$ it consists of the single line

$$\mathcal{C}_1^{+}(0) = \left\{ (\mu,\nu) \in (\mathbb{R}_0^+)^2 : \mu = \left(\frac{\pi_q}{2L}\right)^q \right\}.$$

Again, we can explicitly describe the (NPFS) at infinity and at zero. In particular, when $p = q$, the (NPFS) at infinity is $\mathcal{S}^{+}(\infty) = \cup_{i=1}^{\infty} \mathcal{C}_i^{-}(\infty)$, where

$$\mathcal{C}_{2j-1}^{-}(\infty) = \left\{ (\mu,\nu) \in (\mathbb{R}_0^+)^2 : \frac{j-1}{\mu^{\frac{1}{p}}} + \frac{j}{\nu^{\frac{1}{p}}} = \frac{L}{\pi_p} \right\}$$

and

$$\mathcal{C}^{-}_{2j}(\infty) = \left\{(\mu,\nu) \in (\mathbb{R}^{+}_{0})^{2} : \frac{j}{\mu^{\frac{1}{p}}} + \frac{j}{\nu^{\frac{1}{p}}} = \frac{L}{\pi_p}\right\}.$$

To show the role of the strong asymmetry when $p \neq q$ for the map ϕ defined in (8), we consider f satisfying (f_1) and assume that $\frac{f(s)}{\phi(s)} \to a^{\pm}$ as $s \to 0^{\pm}$ and $\frac{f(s)}{\phi(s)} \to A^{\pm}$ as $s \to \pm\infty$, for some positive constants $a^{\pm}$ and $A^{\pm}$. Then, if $p > q$, it follows that problem (P) has at least one solution with positive slope at $t = 0$ if $A^{+} > \left(\frac{\pi_p}{2L}\right)^{p}$ and it has at least one solution with negative slope at $t = 0$ if $a^{-} > \left(\frac{\pi_q}{2L}\right)^{q}$.

We remark that, according to [20], the same spectra at infinity (or, respectively, at zero) is obtained, and therefore, the same application to problem (P) will occur for any function ϕ having the form of

$$\phi(s) = \begin{cases} \psi_1(s), & \text{for } s \geq 0, \\ \psi_2(s), & \text{for } s \leq 0, \end{cases}$$

where $\psi_1 : \mathbb{R}^{+} \to \mathbb{R}^{+}$ and $\psi_2 : \mathbb{R}^{-} \to \mathbb{R}^{-}$ are increasing bijections with $\psi_i(0) = 0$ and

$$\lim_{s \to +\infty,\, 0^{+}} \frac{\psi_1(\sigma s)}{\psi_1(s)} = \sigma^{p-1}, \quad \text{for all } \sigma > 0,$$

and

$$\lim_{s \to -\infty,\, 0^{-}} \frac{\psi_2(\sigma s)}{\psi_2(s)} = \sigma^{q-1}, \quad \text{for all } \sigma > 0,$$

for some $p, q > 1$.

Example 2. Continuing along this direction, we could consider (like it was done in [20] for an odd ϕ), the case of a ϕ-function such that for some $p, q > 1$,

$$\lim_{s \to \pm\infty} \frac{\phi(\sigma s)}{\phi(s)} = \sigma^{p-1}, \quad \text{for all } \sigma > 0, \tag{9}$$

and

$$\lim_{s \to 0^{\pm}} \frac{\phi(\sigma s)}{\phi(s)} = \sigma^{q-1}, \quad \text{for all } \sigma > 0. \tag{10}$$

Hence, we have a (PPFS) and also a (NPFS) at infinity and one at zero, which are the same like those for the p-Laplacian and the q-Laplacian, respectively. Observe that here we do not assume the ϕ to be odd. For instance, a map of the form

$$\phi : s \mapsto \begin{cases} \phi_p(s), & s \leq -1, \\ \phi_q(s), & -1 \leq s \leq 0, \\ \log(1 + s^{q-1}), & 0 \leq s \leq 1, \\ \phi_p(s)\log(1+s), & s \geq 1, \end{cases}$$

is suitable for our applications.

As a consequence of Theorem 2.1, we have the following

COROLLARY 4.1. *Let ϕ satisfy (ϕ_1), (9), and (10). Suppose that there are positive numbers a^+, a^-, A^+, A^- such that*

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = a^+, \quad \lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = a^-$$

and

$$\lim_{s\to +\infty} \frac{f(s)}{\phi(s)} = A^+, \quad \lim_{s\to -\infty} \frac{f(s)}{\phi(s)} = A^-.$$

Assume also that there are $k, \ell \in \mathbb{N}$ with $k \neq \ell$ such that, for $k = 2j$

$$\frac{j}{(a^+)^{\frac{1}{q}}} + \frac{j-1}{(a^-)^{\frac{1}{q}}} < \frac{L}{\pi_q} < \frac{j}{(a^+)^{\frac{1}{q}}} + \frac{j}{(a^-)^{\frac{1}{q}}}$$

and for $k = 2j+1$

$$\frac{j}{(a^+)^{\frac{1}{q}}} + \frac{j}{(a^-)^{\frac{1}{q}}} < \frac{L}{\pi_q} < \frac{j+1}{(a^+)^{\frac{1}{q}}} + \frac{j}{(a^-)^{\frac{1}{q}}}$$

while, for $\ell = 2i$,

$$\frac{i}{(A^+)^{\frac{1}{p}}} + \frac{i-1}{(A^-)^{\frac{1}{p}}} < \frac{L}{\pi_p} < \frac{i}{(A^+)^{\frac{1}{p}}} + \frac{i}{(A^-)^{\frac{1}{p}}}$$

and for $\ell = 2i+1$,

$$\frac{i}{(A^+)^{\frac{1}{p}}} + \frac{i}{(A^-)^{\frac{1}{p}}} < \frac{L}{\pi_p} < \frac{i+1}{(A^+)^{\frac{1}{p}}} + \frac{i}{(A^-)^{\frac{1}{p}}}$$

Then, problem (P) has at least $|k-\ell|$ solutions with $u'(0) > 0$.

Clearly, a symmetric result holds for the (NPFS).

To show an application of Corollary 4.1 which resembles the one given in the Introduction (this time for $q = 2$), we consider the following situation.

Example 3. Let $\phi : \mathbb{R} \to \mathbb{R}$ be an increasing bijection which is of class C^1 in a neighborhood of zero, with $\phi'(0) > 0$ and satisfies (9) for some $p > 1$. Suppose also that we have

$$\lim_{s\to 0^+} \frac{f(s)}{\phi(s)} = \lim_{s\to +\infty} \frac{f(s)}{\phi(s)} = a$$

$$\lim_{s\to 0^-} \frac{f(s)}{\phi(s)} = \lim_{s\to -\infty} \frac{f(s)}{\phi(s)} = b$$

Assume that (a,b) is in some region $Z^+_{j_1}(0) \cap Z^-_{j_2}(0)$. Then, for $p > 1$ sufficiently large, problem (P) for the interval $[0,\pi]$ has at least $j_1 - 2$ solutions with $u'(0) > 0$ and at least $j_2 - 2$ solutions with $u'(0) < 0$.

Indeed, $L/\pi_p = \pi/\pi_p \to \pi/2 < 2$ as $p \to +\infty$ and, at the same time, $a^{-\frac{1}{p}} + b^{-\frac{1}{p}} \to 2$, so that the pair (a,b) belongs to the second region for both the positive and the negative (PFS) at infinity, for large p.

We remark that, in any case, $|j_1 - j_2| \le 1$.

An example of a function like the ϕ considered here is the following:

$$\phi(s) = \frac{\log(1+|s|)}{\log(1+|s|^{1-p})}\,\mathrm{sgn}(s)\,.$$

Appendix

We present here some technical estimates for the comparison of the time-mappings associated to the quasilinear differential equations

$$(\phi(u'))' + g_1(u) = 0 \tag{A.1}$$

and

$$(\phi(u'))' + g_2(u) = 0 \tag{A.2}$$

where, throughout this section, we assume that $\phi : \mathbb{R} \to \mathbb{R}$ is an increasing bijection with $\phi(0) = 0$ satisfying (ϕ_1) and $g_1, g_2 : \mathbb{R} \to \mathbb{R}$ are continuous functions satisfying assumption (f_1). We also denote by $G_i(s) = \int_0^s g_i(\xi)\,d\xi$ the primitives of g_i.

According to the notation previously introduced, we consider the time-mappings

$$\mathrm{T}_{g_i}(R) := \left| \int_0^R \frac{ds}{\mathcal{L}_r^{-1}(G_i(R) - G_i(s))} \right| + \left| \int_0^R \frac{ds}{\mathcal{L}_l^{-1}(G_i(R) - G_i(s))} \right|$$

(for $i = 1,2$), which represent the distance of two consecutive zeros of a solution of $(\phi(u'))' + g_i(u) = 0$ in an interval where such a solution is positive (negative) and achieves its maximum (minimum) value R.

Lemma A.1. *Assume that there is $R_* > 0$ such that*

$$|g_1(s)| \le |g_2(s)| \quad (\textit{or}\ \ |g_1(s)| < |g_2(s)|\,),$$

holds for all $0 < s \le R_*$, or, respectively, for all $-R_* \le s < 0$. Then, $\mathrm{T}_{g_2}(R) \le \mathrm{T}_{g_1}(R)$ (or $\mathrm{T}_{g_2}(R) < \mathrm{T}_{g_1}(R)$), for all $0 < R \le R_*$ or, respectively, for all $-R_* \le R < 0$.

Proof. We consider only the case when $R > 0$, being the other one completely symmetric. Then, by the assumption, given any $R \in]0, R_*]$, we have that $g_1(s) \le g_2(s)$ (or $g_1(s) < g_2(s)$) for all $s \in]0, R]$. Hence,

$$G_1(R) - G_1(s) = \int_s^R g_1(\xi)\, d\xi \le \int_s^R g_2(\xi)\, d\xi = G_2(R) - G_2(s)$$

(or $G_1(R) - G_1(s) < G_2(R) - G_2(s)$) holds for each $0 < s \le R$. From this and the definition of $\mathrm{T}_{g_i}(R)$, the result immediately follows. □

LEMMA A.2. *Assume that there is $R^* > 0$ such that*

$$|g_1(s)| \le |g_2(s)|,$$

holds for all $s \ge R^*$, or, respectively, for all $s \le -R^*$. Then, for each $\varepsilon > 0$ there is $R_\varepsilon > R^*$ such that

$$\mathrm{T}_{g_2}(R) \le \mathrm{T}_{g_1}(R) + \varepsilon$$

holds for all $R > R_\varepsilon$, or, respectively, for all $R < -R_\varepsilon$.

Proof. As before, we discuss only the case when $R > 0$. Let u_i be the solution of $(\phi(u'))' + g_i(u) = 0$, for $i = 1, 2$, with $u_i(0) = 0$ and $\max u_i = u(t_i^*) = R > R^*$. From the equation, we see that u_i is strictly increasing in $[0, t_i^*]$ and strictly decreasing in $[t_i^*, \mathrm{T}_{g_1}(R)]$. Hence there are uniquely determined t_i^- and t_i^+, with $0 < t_i^- < t_i^* < t_i^+ < \mathrm{T}_{g_1}(R)$, such that $u_i(t) \ge R^*$ for $t \in [0, \mathrm{T}_{g_1}(R)]$ if and only if $t \in [t_i^-, t_i^+]$. By the same argument like in the proof of Lemma A.1, it easily follows that

$$\begin{aligned}\int_{R^*}^R \frac{ds}{\mathcal{L}_r^{-1}(G_2(R) - G_2(s))} &+ \int_{R^*}^R \frac{ds}{|\mathcal{L}_l^{-1}(G_2(R) - G_2(s))|} \\ &\le \int_{R^*}^R \frac{ds}{\mathcal{L}_r^{-1}(G_1(R) - G_1(s))} + \int_{R^*}^R \frac{ds}{|\mathcal{L}_l^{-1}(G_1(R) - G_1(s))|}.\end{aligned}$$

Therefore, we have that

$$\begin{aligned}&\mathrm{T}_{g_2}(R) - \mathrm{T}_{g_1}(R) \\ &\quad\le \sum_{i=1,2} \left((t_i^+ - t^*) + (t^* - t_i^-)\right) \\ &\quad= \sum_{i=1,2} \left(\int_0^{R^*} \frac{ds}{\mathcal{L}_r^{-1}(G_i(R) - G_i(s))} + \int_0^{R^*} \frac{ds}{|\mathcal{L}_l^{-1}(G_i(R) - G_i(s))|} \right) \\ &\quad\le \sum_{i=1,2} \left(\frac{R^*}{\mathcal{L}_r^{-1}(G_i(R) - G_i(R^*))} + \frac{R^*}{|\mathcal{L}_l^{-1}(G_i(R) - G_i(s))|} \right)\end{aligned}$$

holds for each $R > R^*$. At this point, the result easily follows, using (f_1) and letting $R \to +\infty$. □

The next result is a straightforward consequence of Lemma A.2

COROLLARY A.1. *Assume that there is $R^* > 0$ such that*

$$|g_1(s)| \leq |g_2(s)|,$$

holds for all $s \geq R^*$, or, respectively, for all $s \leq -R^*$. Then

$$\limsup_{R\to+\infty} \mathrm{T}_{g_2}(R) \leq \liminf_{R\to+\infty} \mathrm{T}_{g_1}(R)$$

(respectively, $\limsup_{R\to-\infty} \mathrm{T}_{g_2}(R) \leq \liminf_{R\to-\infty} \mathrm{T}_{g_1}(R)$*).*

REFERENCES

[1] M. ABRAMOWITZ AND I. STEGUN, *Handbook of mathematical functions*, Dover, New York (1965).

[2] L. BOCCARDO, P. DRÁBEK, D. GIACHETTI AND M. KUČERA, *Generalization of Fredholm alternative for nonlinear differential operators*, Nonlinear Anal. **10** (1986), 1083–1103.

[3] A. CAPIETTO, *An existence result for a two-point superlinear boundary value problem*, Proceedings of the 2nd International Conference on Differential Equations, Marrakesh (1995).

[4] A. CAPIETTO AND W. DAMBROSIO, *Multiplicity results for some two-point superlinear asymmetric boundary value problem*, Nonlinear Anal. **38** (1999), 869–896.

[5] A. CAPIETTO AND W. DAMBROSIO, *Boundary value problems with sublinear conditions near zero*, NoDEA Nonlinear Differential Equations Appl. **6** (1999), 149–172.

[6] A. CAPIETTO, J. MAWHIN AND F. ZANOLIN, *Boundary value problems for forced superlinear second order ordinary differential equations*, in H. BREZIS, *Nonlinear partial differential equations and their applications*, Taylor & Francis (1994).

[7] A. CAPIETTO, J. MAWHIN AND F. ZANOLIN, *On the existence of two solutions with a prescribed number of zeros for a superlinear two-point boundary value problem*, Topol. Methods Nonlinear Anal. **6** (1995), 175–188.

[8] W. DAMBROSIO, *Time-map techniques for some boundary value problems*, Rocky Mountain J. Math. **28** (1998), 885–926.

[9] W. DAMBROSIO, *Multiple solutions of weakly-coupled systems with p-Laplacian operators*, Results Math. **36** (1998), 34–54.

[10] M. DEL PINO, M. ELGUETA AND R. MANÁSEVICH, *A homotopic deformation along p of a Leray-Schauder degree result and existence for* $(|u'|^{p-2}u')' + f(t,u) = 0$, $u(0) = u(T) = 0$, $p > 1$, J. Differential Equations **80** (1992), 1–13.

[11] M.A. DEL PINO, R.F. MANÁSEVICH AND A.E. MURÚA, *Existence and multiplicity of solutions with prescribed period for a second order quasilinear ODE*, Nonlinear Anal. **18** (1992), 79–92.

[12] G. Dinca and L. Sanchez, *Multiple solutions of boundary value problems: an elementary approach via the shooting method*, NoDEA Nonlinear Differential Equations Appl. **1** (1994), 163–178.

[13] P. Drábek, *Solvability of boundary value problems with homogeneous ordinary differential operator*, Rend. Istit. Mat. Univ. Trieste **18** (1986), 105–124.

[14] P. Drábek and R. Manásevich, *On the closed solution to some nonhomogeneous eigenvalue problems with p-Laplacian*, Differential Integral Equations **12** (1999), 773–788.

[15] S. Fučík, *Solvability of nonlinear equations and boundary value problems*, Reidel, Dordrecht (1980).

[16] M. García-Huidobro, R. Manásevich and F. Zanolin, *Strongly nonlinear second order ODE's with unilateral conditions*, Differential Integral Equations **6** (1993), 1057–1078.

[17] M. García-Huidobro, R. Manásevich and F. Zanolin, *A Fredholm-like result for strongly nonlinear second order ODE's*, J. Differential Equations **114** (1994), 132–167.

[18] M. García-Huidobro, R. Manásevich and F. Zanolin, *On a pseudo Fučík spectrum for strongly nonlinear second order ODE's and an existence result*, J. Comput. Appl. Math. **52** (1994), 219–239.

[19] M. García-Huidobro, R. Manásevich and F. Zanolin, *Infinitely many solutions for a Dirichlet problem with a nonhomogeneous p-Laplacian-like operator in a ball*, Adv. Differential Equations **2** (1997), 203–230.

[20] M. García-Huidobro and P. Ubilla, *Multiplicity of solutions for a class of second-order equations*, Nonlinear Anal. **28** (1997), 1509–1520.

[21] P. Habets, M. Ramos and L. Sanchez, *Jumping nonlinearities for Neumann BVPs with positive forcing*, Nonlinear Anal. **20** (1993), 533–549.

[22] M. Henrard, *Infinitely many solutions of weakly coupled superlinear systems*, Adv. Differential Equations **2** (1997), 753–778.

[23] R. Manásevich and F. Zanolin, *Time-mappings and multiplicity of solutions for the one-dimensional p-Laplacian*, Nonlinear Anal. **21** (1993), 269–291.

[24] Z. Opial, *Sur les périodes des solutiones de l'équation différentielle* $x'' + g(x) = 0$, Ann. Polon. Math. **10** (1961), 49–72.

[25] C. Rebelo, *A note on uniqueness of Cauchy problems associated to planar Hamiltonian systems*, Portugal. Math. **57** (2000), 415–419.

Authors' addresses:

Marta García-Huidobro
Departamento de Matemáticas
P. Universidad Católica de Chile
Casilla 306, Correo 22, Santiago, Chile
E-mail: `mgarcia@mat.puc.cl`

Raúl Manásevich
Centro de Modelamiento Matemático and Departamento de Ingeniería Matemática
Universidad de Chile
Casilla 170, Correo 3, Santiago, Chile
E-mail: manasevi@dim.uchile.cl

Fabio Zanolin
Dipartimento di Matematica e Informatica
Università di Udine
Via delle Scienze 206, 33100 Udine, Italy
E-mail: zanolin@dimi.uniud.it

Received November 16, 2011

Editorial Note

The Managing Editors would like to remind the authors of the ethical rules set out in the document of IMU "Best current practices for Journals" (see www.mathunion.org). In particular, we stress the following:

"Authors must abide by high standards of research integrity and good scholarship. It is the responsibility of authors to submit a well written, mathematically correct article, if necessary seeking advice if it is not written in their native language, to clearly describe any novel and non-trivial content, and to suitably acknowledge the contributions of others, including referees. Submission of a paper to a journal implies that it is not currently under consideration by any other journal, and that any substantial overlap with other published or submitted papers is duly acknowledged. In addition authors should be responsive to correspondence with the journal. Multiple authors should communicate fully, speak with one voice, and accept mutual responsibility in their communications with the journal. All authors are expected to have materially contributed to the paper, and to be familiar with its content."

Alessandro Fonda
Emilia Mezzetti
Pierpaolo Omari
Maura Ughi

Contents

Editorial Policy

The journal *Rendiconti dell'Istituto di Matematica dell'Università di Trieste* publishes original articles in all areas of mathematics. Special regard is given to research papers, but attractive expository papers may also be considered for publication.

The journal usually appears in one issue per year. Additional issues may however be published. In particular, the Managing Editors may consider the publication of supplementary volumes related to some special events, like conferences, workshops, and advanced schools.

All submitted papers will be refereed. Manuscripts are accepted for review with the understanding that the work has not been published before and is not under consideration for publication elsewhere.

The journal can be obtained by exchange agreements with other similar journals.

Instructions for Authors

Authors are invited to submit their papers by e-mail directly to one of the Managing Editors in PDF format.

All the correspondence regarding the submission and the editorial process of the paper are done by e-mail.

Papers have to be written in one of the following languages: English, French, German, or Italian. Abstracts should not exceed ten printed lines, except for papers written in French, German, or Italian, for which an extended English summary is required.

After acceptance, manuscripts have to be prepared in LaTeX using the style *rendiconti.cls* which can be downloaded from the web page.

Any figure should be recorded in a single PDF, PS (PostScript), or EPS (Encapsulated PostScript) file.

INDIRIZZO DEL COMITATO DI REDAZIONE

Rendiconti dell'Istituto di Matematica dell'Università di Trieste
Dipartimento di Matematica e Informatica - Università - Via Valerio 12/1 - 34127 Trieste (Italy)
Tel. +39-040-5582635 — Fax +39-040-5582636 — Pagina web: www.dmi.units.it/~rimut/

Direttore Responsabile: DANIELE DEL SANTO

Periodico registrato il 25 Settembre 1968 al n. 358 del registro dei periodici del Tribunale di Trieste

Annual subscription rate: €50,00 (one issue, inclusive of any supplements).
Subscription inquiries should be sent to
EUT, Edizioni Università di Trieste, Piazzale Europa 1, 34127 Trieste, Italia.

ISSN 0049-4704

Contents

ISSN 0049-4704